Strategies for continuous improvement and improved competitiveness for the sustainable bio-based industries

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Christoph Metzner

December 2018

ACKNOWLEDGEMENTS

Funding for this project was provided by the U.S. Department of Energy research grant as administered by The University of Tennessee R11-3215-096 and the United States Department of Agriculture (USDA) Forest Service and McIntire-Stennis TENOOMS-107 administered by The University of Tennessee Agricultural Experiment Station. Additional funding was also provided by the U.S. Forest Service, Forest Products Laboratory in Madison, Wisconsin under research grant R11-2219-690 as administered by The University of Tennessee.

I would like to express my sincere gratitude to Dr. Timothy M. Young for his guidance, support, and encouragement through all my tasks here at the University of Tennessee. Additionally, Dr. Young thank you for providing me the possibility to study at the University of Tennessee.

Furthermore, I would also like to express my gratitude to Dr. Alexander Petutschnigg and Dr. Marius-Catalin Barbu for their knowledge and support from the Salzburg University of Applied Sciences throughout the past year.

I would like to thank my committee Dr. Bogdan Bichescu and Dr. Terry Liles for their support and suggestions throughout my studies and creation of my thesis.

Special thanks to the coworkers from the Center for Renewable Carbon, Mr. Chris Helton, Mr. Anton Astner, and Mr. Qijun Zhang for your support and companionship. Furthermore, I would like to express my sincere gratitude to Ms. Wendy Garcia for her support and companionship throughout my year here in the United States.

Finally, I would like to thank my family for their incredible support throughout all my studies and tasks over the past years.

ABSTRACT

Cellulosic biomass is a highly variable feedstock. The large variation in key quality attributes (*e.g.,* ash content, moisture content, and particle size) challenges the consistency of the feedstock supply from a technological and economical perspective. This affects the cost and the overall competitiveness of the sustainable bio-based industries. This research focuses on developing strategies to reduce variation and cost throughout the supply chain for the bio-based industries.

The goal of this research is to provide practitioners with tools to quantify variation of the components of the supply chain and illustrate that variation accumulates throughout the supply chain which induces costs from higher than necessary operational targets. The objectives of this research are: 1) develop quality loss functions for the components of the biomass supply chain; 2) create a simulation model suitable to quantify feedstock variation; 3) characterize the impact of variation on the financial loss, and 4) develop a handbook of statistical and continuous improvement techniques to promote variation reduction.

The Excel simulation model uses Statistical Process Control and Taguchi's Loss Function combined with Galton's theory of 'components of variance' to estimate the financial loss due to variation. Sensitivity analyses are used to characterize the impact of variation on loss for ash content, moisture content, and particle size. The handbook provides practitioners with a guide for improved application of universally accepted key continuous improvement techniques.

The additional loss per unit on average for Switchgrass from ash content variation was estimated to be \$17.33 per dry ton, while for particle size (woody residues) the loss was \$10.32 per dry ton. The additional loss per unit on average for moisture content variation was estimated for an example supply chain. The loss per unit for harvest/collection was \$2.02, transport was \$4.93, drying was \$3.19, and densification was \$13.23 per dry ton. The results of this study suggest that Taguchi's Loss Functions are suitable to estimate the loss for feedstock quality characteristics based on variation. The simulation tool and handbook will help practitioners of the sustainable bio-based industries improve the supply chain's performance (available at www.spc4lean.com).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ATTACHMENTS

CHAPTER ONE
INTRODUCTION

In prior decades the sustainable bio-based industries have faced major technological and economic challenges. For example, the cellulosic biofuel industry had to develop efficient conversion technologies and supply chain systems to produce economic viable biofuels. These biofuels had to be produced with non-edible cellulosic biomass at a cost that is competitive with conventional fuels (Yue et al., 2013). Another example, is the forest products industry which was heavily impacted by the collapse of the United States housing market during the economic crisis from 2007 to 2009 (Howard and Jones, 2016). Today, both industries face competitive pressures through increased globalization and procuring cost-competitive raw material supply, *e.g., large feedstock variations induce variation in the process and final product.* The large feedstock variations lead to increased costs, *i.e., higher than necessary operational targets for weight, solvents, resin, etc. must be maintained given the large variations in raw materials* (Kenney et al., 2013, Salim and Johansson, 2016).

For example, based on the literature the supply chain costs for producing cellulosic ethanol (*i.e.,* biofuel) are roughly 35% of the total production costs (Hess et al., 2007, You et al., 2012). Given the current production costs for cellulosic ethanol $5.90 (ranging between $5.06 to $6.73/GGE) (Warner et al., 2017), based on a gasoline gallon equivalent (GGE), the sole supply chain costs would equivalate to $2.07/GGE (ranging between $1.77 to $2.32/GGE) (Table 1). These supply chain costs represent already 84% of the total production costs for corn-grain based ethanol of $2.46/GGE (ranging between $1.50/GGE to $4.56/GGE) (ISU, 2018). However, both types of ethanol cannot currently compete with the crude oil price of $1.62/gallon (*i.e.,* $68/barrel) (Macrotrends LLC, 2018a), which is reflected by the historic U.S. retail price for gasoline and ethanol (*i.e.,* E85) (Figure 1). For cellulosic ethanol to be competitive against wholesale gasoline prices, achieved with crude oil of $100 per barrel, the production cost of cellulosic ethanol must be $3 per gallon (Sims et al., 2010). Given the current crude oil prices this number must be reduced even further.

Table 1. Production cost comparison of various fuel types.

Fuel type	Cellulosic Ethanol	Corn-Grain Ethanol	Crude Oil
Year data is from	~ 2015	2007-2018	2007-2018 (July 2018)
Production costs	\$5.90/GGE[1]	\$2.46/GGE	\$1.85/Gallon (\$1.62/Gallon)
Range	\$5.06-\$6.73/GGE	\$1.50-\$4.56/GGE	\$0.97-\$2.61/Gallon
Reference	(Warner et al., 2017)	(ISU, 2018)	(Macrotrends LLC, 2018a)

[1] GGE is the amount of fuel it takes to equal the energy content of one liquid gallon of gasoline.

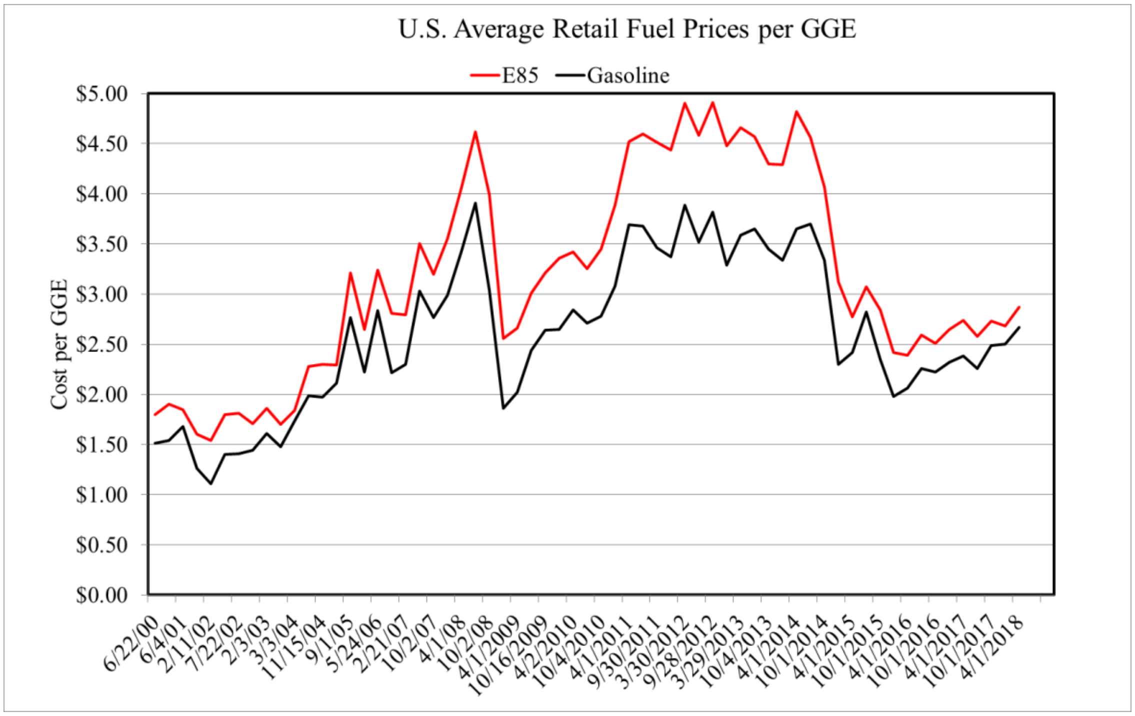


Figure 1. U.S. average retail fuel prices per GGE for gasoline and ethanol (E85) (DOE, 2018).

Thus, to achieve price competitive products, companies of the sustainable bio-based industries must rely on analytics and statistical methods to quantify variation of key input variables in their production systems (or supply chain systems). Methodologies such as statistical process control, lean or the Toyota Production System (TPS), etc., are presented in this thesis as examples of suitable techniques to improve processes.

Rationale and Thesis Execution

A main problem for these sustainable bio-based industries is the cost-efficient supply of the highly variable raw materials. This high variability in key quality characteristics challenges the performance of each component of the supply chain and manufacturing system (Germain et al., 2008). Raw material variation and the occurring variability in process execution influences the final quality of the product (Sofuoglu and Kurtoglu, 2012). Therefore, companies with highly variable product attributes (*e.g.*, density, strength, yield, etc.) are less competitive in the market as enterprises producing items with little variation. As a result, most manufacturers must compensate for excessive raw material variation with higher targets in their key process variables (*e.g.*, weight, resin, etc.) to meet final product specification, which ultimately lead to increased costs (Taguchi et al., 2004). These unnecessary costs through raw material variation exacerbate the already tense economic position of the cellulosic biomass supply chain within the total biofuel production costs. Reducing process or supply variation is desirable since operational targets can be decreased which improves financial performance. Through the correct application of continuous improvement techniques engineers and managers can identify sources of variation which facilitates efforts to reduce variation in the manufacturing process (or supply chain).

Previous studies emphasize the use of statistical process control (SPC) to improve performance of production or supply chain processes. For example, the application of real-time control charts has improved performance of many forest product manufacturers (André and Young, 2013, Astner et al., 2015, Carty et al., 2015, Maness et al., 2003, Riegler et al., 2015, Steiner et al., 2017, Young and Winistofer, 1999, Young et al., 2007, Young et al., 2014, Young et al., 2015a, Young et al., 2015b). This research expands upon earlier research where a simulation model for quantifying variation in the *'bio-depot'* concept for the biofuel industry was developed (Platzer, 2016).

This study enhances the previous research from Platzer (2016) by developing more strategies and techniques to improve the biomass supply chain to enhance the competitiveness of products from the sustainable bio-based industries by lowering costs. A more advanced model to simulate the financial loss using the Taguchi Loss Function combined with Galton's theory of components of variance for estimating financial loss due to variation in the feedstock supply chain system was developed as part of this thesis. Variation was simulated from some existing data and enhanced with bootstrapping. Ash content, moisture content, and particle size were the variables in the supply chain that were modeled.

The simulation model is intended to help practitioners identify the components of the system with the largest variations and highest costs. Statistical process control (SPC) procedures and Taguchi's quality loss functions were used in the model to improve the visualization and quantification of variation that occurs throughout the supply chain system. This improved visualization is achieved through graphical display of the variation and loss. The continuous improvement techniques used in the thesis were summarized into a handbook for practitioners to improve the application of these helpful and universally accepted techniques for promoting variation reduction and cost savings.

Hypothesis, Goal, and Objectives

The research hypothesis aims to determine whether continuous improvement techniques are suitable to quantify variation of raw material quality characteristics affecting supply chain and costs. The goal of this research is to provide practitioners of the sustainable bio-based industries with tools to quantify variation of the components of the supply chain and illustrate that variation accumulates throughout the supply chain which induces cost. Based on the goal of this thesis, the following objectives were formulated:

- Development of quality loss functions to quantify the monetary loss through feedstock variation across the supply chain and its components;
- Development of a simulation-tool for practical application of these quality loss functions;
- Conduct sensitivity analyses to characterize the impact of variation on the loss computed with the developed loss functions;
- Development of a continuous improvement handbook for the sustainable bio-based industries.

A brief introduction of the cellulosic biofuel industry and forest products industry is presented. The biofuel industry can be classified into unprocessed (*e.g.,* pellets or firewood) and processed (*e.g.,* charcoal, ethanol, or biogas) biofuels (FAO, 2008); in context of the thesis the second class is referred as biofuel industry. The wood product industry, such as producers of furniture, wood composites, engineered wood panels, and construction timber, etc., is referred to as the forest products industry in this thesis.

Biofuels Industry

Rising global energy demand with corresponding limited reserves of conventional energy sources has created a renewed focus on alternative energy policies. Even though current energy prices for oil and natural gas are at much lower levels than ten years ago (Figure 2), scientists and governments are still engaged in the development of policies and technologies for alternative energy generation (Guo et al., 2015). Using biomass as a renewable energy source, next to solar, wind, or water, has promise as noted by Gold and Seuring (2011). Bioenergy is created from different types of biomass and can be a viable substitute for conventional fossil fuels (Gold and Seuring, 2011). Studies have indicated the positive effects of producing biofuels for the United States, *e.g., ensuring energy security by reducing dependency on foreign petroleum imports, economic development for rural communities, and mitigation of greenhouse gases* (Ekşioğlu et al., 2009, Mabee et al., 2011).

Initially, biofuels were produced from sugar-based feedstocks such as corn and sugarcane. Unfortunately, despite having great benefits, using edible feedstocks to produce biofuels sparked a heated discussion in the population about the optimal usage, *i.e., using edible biomass as fuels instead of food considering the scarcity of food worldwide.* For example, the use of corn for biofuel production increased the prices of food commodities (Tyner, 2010). To overcome these challenges renewable fuel standards across the globe were introduced to promote the production of biofuels using non-edible biomass feedstocks.

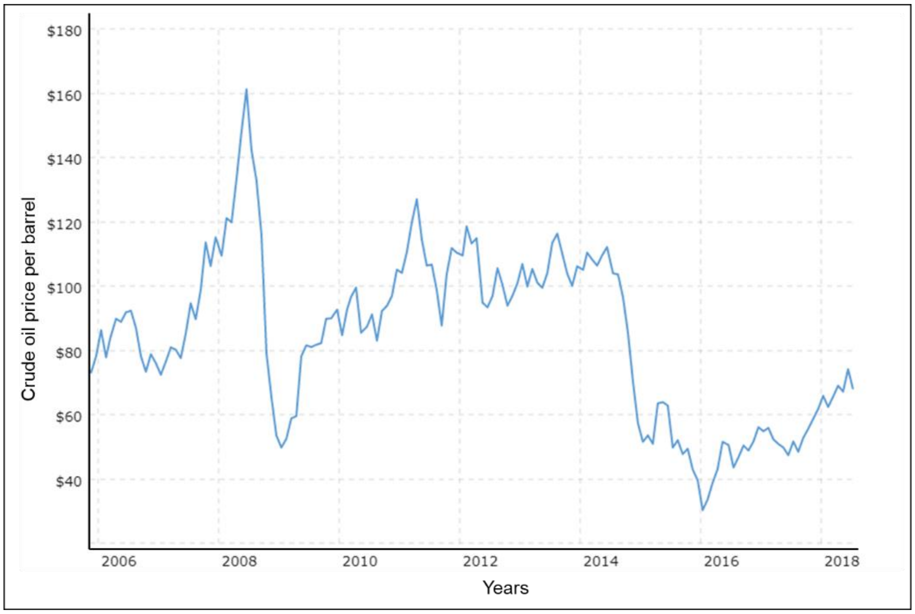


Figure 2. West Texas Intermediate (WTI) crude oil ‘real prices’ per barrel developments (Macrotrends LLC, 2018a).

The *Energy Independence and Security Act of 2007* was passed by the U.S. congress which mandates that by the year 2022 at least 36 billion gallons of biofuel (*e.g.,* ethanol or biodiesel) are produced annually (EISA, 2007). To promote the production of cellulosic ethanol only 15 billion gallons of biofuel can stem from edible biomass. This policy pushed research efforts into developing efficient conversion technologies, pretreatment methods, and efficient supply chain systems for lignocellulosic feedstocks (Daoutidis et al., 2013). The advantages of lignocellulosic feedstocks lie in their abundant occurrence in the United States, the lack of already established customer markets, as well as not competing against food crops for traditional production land (Hoekman, 2009). Despite these benefits technological and logistical challenges remain mostly through the high variability of the feedstock quality which significantly impacts the yield of biofuel production (Kenney et al., 2013). This variation in feedstock quality characteristics affects all components of the supply chain and conversion processes and increases costs. For example, depending on the feedstock type the harvesting window is seasonal, which makes it necessary to store the biomass, however storage may increase moisture content resulting in higher material degradation (Lamers et al., 2015). Furthermore, lignocellulosic biomass has lower bulk density, which paired with increased moisture content increases transportation costs (Lin et al., 2016). An optimal and sustainable supply of biomass to the conversion facility to maintain stable costs of feedstock supply, which typically account for 20% to 40% of the total production costs of ethanol, is imperative (Angus-Hankin et al., 1995). Thus, modeling supply chain systems which quantify variability of biomass quality (*e.g.,* ash content, moisture content, and particle size studied in this thesis) and estimate costs are vital as a first step in reducing the costs of biofuels; which is the aspiration of this thesis.

Forest Products Industry

The economic state of the forest products industry was characterized by a steady growth with cyclical fluctuations until the end of the last century (Howard and Jones, 2016). Unfortunately, economic turbulences caused uncertainties and change (Nicholls and Bumgardner, 2018) for the industry in the first decade of the 21st century. Economic challenges such as the crisis from 2007 and ongoing globalization of the domestic forest products market aggravated the competition for the U.S. forest product industry (Hansen, 2010). One major problem caused by this internationalization was impairing the price for roundwood and sawn timber. For example, according to Keegan et al. (2011) forestland owners in the Western United States generated higher revenues by exporting their roundwood to Asian customers. As a result, the sales price for roundwood went up and domestic mills had to compete with foreign buyers, which benefit from a different economic background. This unfavorable price structure forced the mill owner to either accept lower margins or greater idle production capacities. As an example, the reduction of employment in the wood product industry by 47% reflected these developments (a reduction from 620,300 jobs in 1999 to 331,000 in 2011) (BLS, 2018). The recovery of the housing market in years after the crisis helped the forest products industry to stabilize (Figure 3). However, key challenges remain such as high raw material prices, increased competition from foreign companies, variability in raw material, and reducing variation in key process variables.

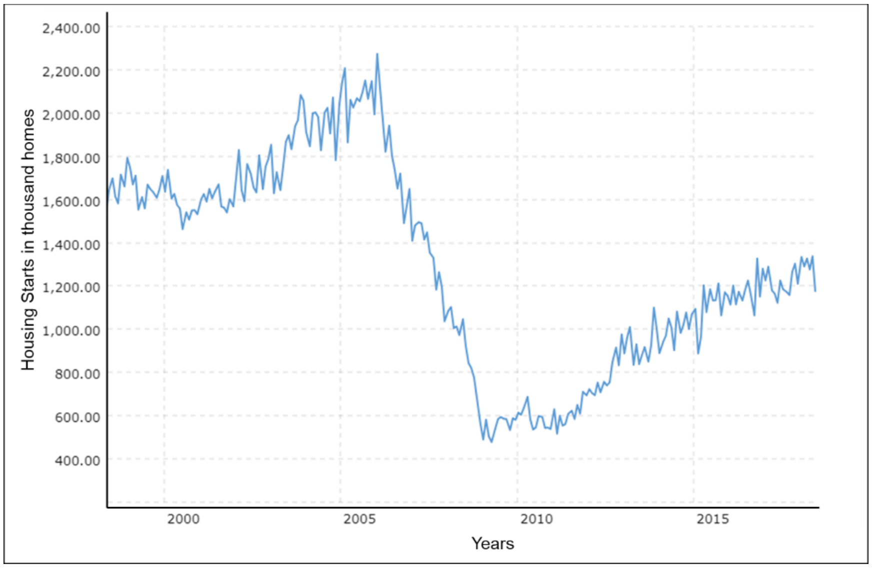


Figure 3. Housing starts in the United States in thousand homes (Macrotrends LLC, 2018b).

This variability in raw material and process variables form one of the greatest challenges for forest products manufacturers, *i.e., to be economically competitive while executing the production process at the lowest cost to generate products with best quality possible* (Salim and Johansson, 2016). The key process variables and the incoming raw material quality determine the performance of each operation in the production chain and the final product quality. Thus, variation in wood has a great impact on the production and generate unnecessary loss (Sofuoglu and Kurtoglu, 2012). Therefore, managers and engineers must develop strategies to increase the efficiency of the production and identify factors that lead to lower costs. Hence, the correct application of continuous improvement techniques, such as statistical process control and lean management, is critical for improvement.

Thesis Organization

The thesis is organized after Chapter One as follows. Chapter Two is a review of the literature on the current state and issues with cellulosic biomass supply chains and methodologies associated with continuous improvement and statistical process control. Chapter Three provides an overview on the materials and methods, and the simulation approach used in this research. Results and discussions are presented as related to the simulation model in Chapter Four. An outline of the continuous improvement handbook for practitioners is given in Chapter Five. Chapter Six is the conclusion and recommendations for future research.

CHAPTER TWO
LITERATURE REVIEW

The literature review presented in this chapter is a general introduction to the methodologies associated with continuous improvement. The intent of the chapter is to provide the underlying framework and justification for the methods used throughout the research study. Given that a vast amount of knowledge exists on this subject matter, and the plethora of literature on the subject, the intent is to provide the reader with a general overview. More detail can be found in the referenced literature.

Biomass Supply Chain

This section provides an overview of the state and design of the biomass supply chain (BSC) for the biorefinery. Various studies discussed the BSC performance and its associated difficulties for individual cases. Alongside the analysis of environmental and social-economic impacts of biofuel production on ambient regions of the biorefinery, mathematical models were used to assess the optimal solution for complex biomass conversion sites and their supply chain systems (Sharma et al., 2013).

The supply chain is an integrated system to process materials into a finished product (Beamon, 1998). Suppliers, manufacturers, distributors, and retailers are the four basic business entities within a supply chain (Beamon, 1998). Whereas, the BSC represents the first two aforementioned entities which consist out of the following components: Feedstock planting and cultivation, harvesting, handling, storage, in-field/forest transportation, road transportation, and preprocessing (Rentizelas et al., 2009). The BSC depends on several aspects but is not limited to feedstock type, region, transport logistics, and biomass conversion technology. A common BSC relies on the “conventional-bale” supply chain design (Figure 4), *e.g., biomass is baled upon harvest and transported as bales to the mill gate.* There are many challenges of this BSC system (Awudu and Zhang, 2012). For example, quantity (or densification) and quality management of harvested biomass, transportation and logistics concerns (*i.e.*, high volume and low weight), and production yields from loss during storage (Awudu and Zhang, 2012).

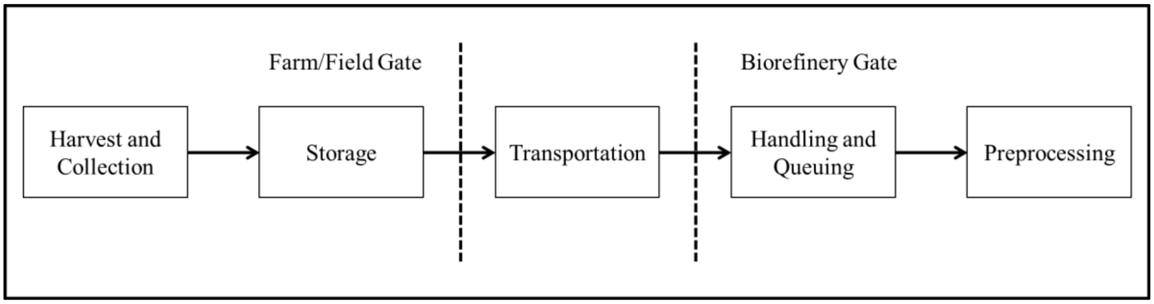


Figure 4. Conventional-bale biomass supply chain for herbaceous lignocellulosic biomass (Jacobson et al., 2014).

Optimal supply chain management for low-bulk density and aerobically unstable biomass is crucial for the performance of biofuel production. However, the "conventional-bale system" requires biorefineries to be located near the supply source, *e.g., within a 50-mile radius* (Argo et al., 2013). Studies from the Idaho National Laboratory showed that these BSC for biorefineries may not meet the rising biofuels production goals due to limited access to proper feedstocks in quantity and quality within a restricted procurement zone (Searcy et al., 2010). Therefore, the advanced uniform-format feedstock supply system (AUD) was developed (Figure 5). The AUD design should reduce some of the aforementioned uncertainties and improve the viability of bioethanol production. The key difference between both designs lies in the positioning of the preprocessing step. Whereas the task of feedstock preprocessing in a conventional design is done by the biorefinery itself; in an AUD this task is positioned immediately after the harvest and collection step (Jacobson et al., 2014). The preprocessing will take place in the so-called *'bio-depot'*, which is closely located to the harvest and collection sites. This allows the production of uniform, aerobically stable, and easy to ship commodity products (Jacobson et al., 2014). Increased collection radius and the liberty of feedstock selection simplifies the process in meeting the specification limits of key feedstock quality characteristics for the specific conversion technology. This practice assures evenly distributed properties, such as ash content, moisture content, and particles size and guarantees steady supply of equal feedstock to the biorefineries (Argo et al., 2013). Typically, the upstream operations in the biomass supply chain (*e.g.*, harvest, preprocessing, etc.) are in control of the final raw material quality. While the financial loss through bad raw material quality is rather experienced at the downstream operations (*e.g.*, transport or biorefinery). Thus, to avoid unnecessary costs all components of the supply chain must collaborate and communicate to guarantee a price competitive end product.

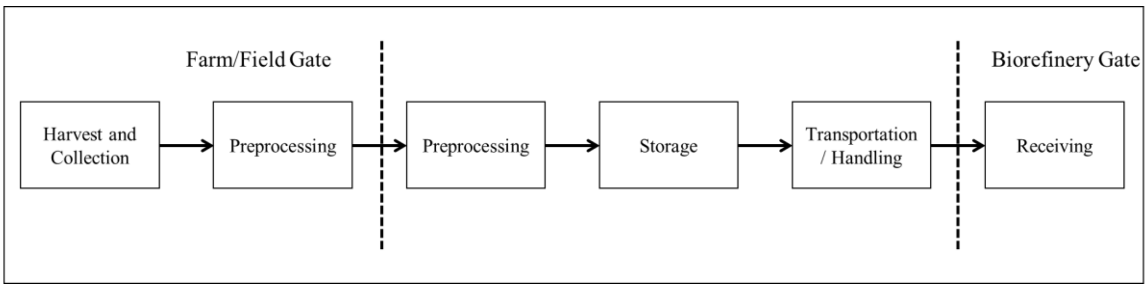


Figure 5. Advanced uniform-format feedstock supply system (AUD) – components (Hess et al., 2009).

The overall performance of the AUD depends on the individual performance of each supply chain component. Harvest and collection of biomass depends on seasonal availability and is energy-intensive; based on machinery used in the harvest / collection operation the biomass can be introduced with contaminates, *e.g., soil* (McKendry, 2002). Unless not immediately processed at the bio-depot or biorefinery seasonal available feedstocks, such as Switchgrass (*Panicum virgatum* L.), must be stored to ensure quality and a stable supply to the biorefinery (Mitchell and Schmer, 2012). For each specific biorefinery supply chain system the type of storage must be selected under economical, qualitative, regional, and feedstock specific aspects (Darr and Shah, 2014). In addition, tarped storage has been found to be an effective way in reducing dry matter loss as well as keeping initial installing costs low at the same time (Darr and Shah, 2014). Transportation costs crucially influence the overall competitiveness of biofuels, *i.e., transportation and handling are non-value-adding operations.* According to Hess et al. (2007) 35% of the production costs stem from feedstock production and logistics, while biomass logistics constitute up to 75% of those costs. Biomass transportation happens either via truck for short or rail for long distances (Figure 6) (Lin et al., 2016).

Depending on the feedstock type biomass in its uncompressed form has a low bulk density of 50 to 130 kg/m³, whereas pellets have a bulk density up to 700 kg/m³ (Sokhansanj and Turhollow, 2004).

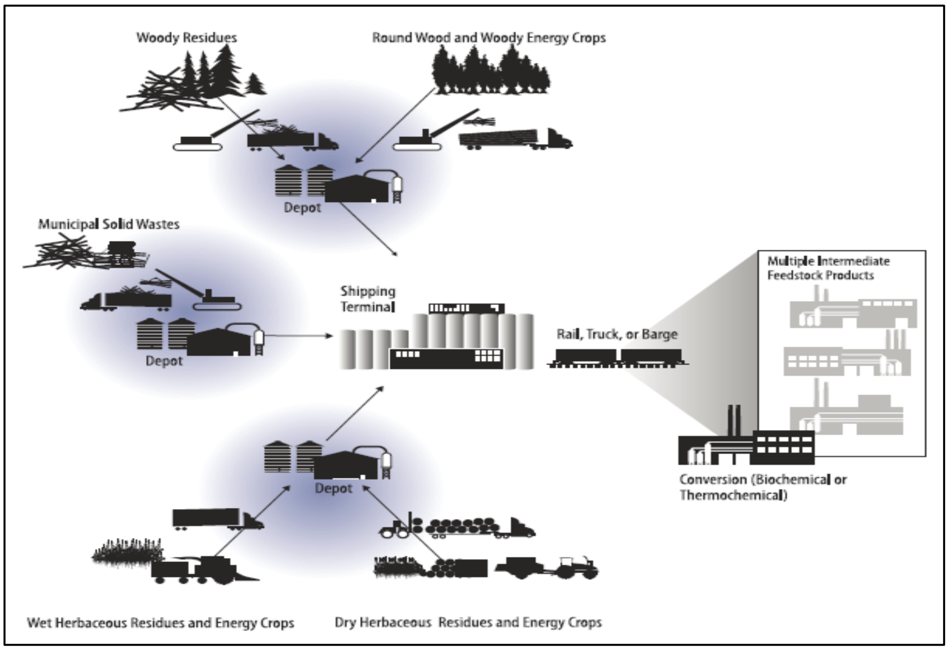


Figure 6. Advanced uniform-format feedstock supply-system (AUD) (Hess et al., 2009).

Low-density materials have higher transportation costs due to volume restrictions of truck trailers. Densified feedstocks are more efficient to handle, however this efficiency is offset by an additional cost step of preprocessing (Lin et al., 2016). Biomass preprocessing significantly increases the potential output of industrial biofuel production sites (Lin et al., 2013). Comminution, drying, blending, and densification are the major operations of a bio-depot supply chain concept (Figure 7), see (Platzer, 2016).

Mechanical particle size reduction – comminution – crucially impacts the biomass conversion process (Marino et al., 2017). Hammer mills are usually used to reduce the size of herbaceous biomass to < ~ 2.5 cm. Also, the initial feedstock moisture content impacts the particle size distribution, grinding energy, and throughput of the hammer mill

(Tumuluru et al., 2016). Comminution is the most cost-intensive operation of the biomass conversion process (Tumuluru et al., 2016). Biomass with high moisture content is usually dried to decrease the grinding energy consumption.

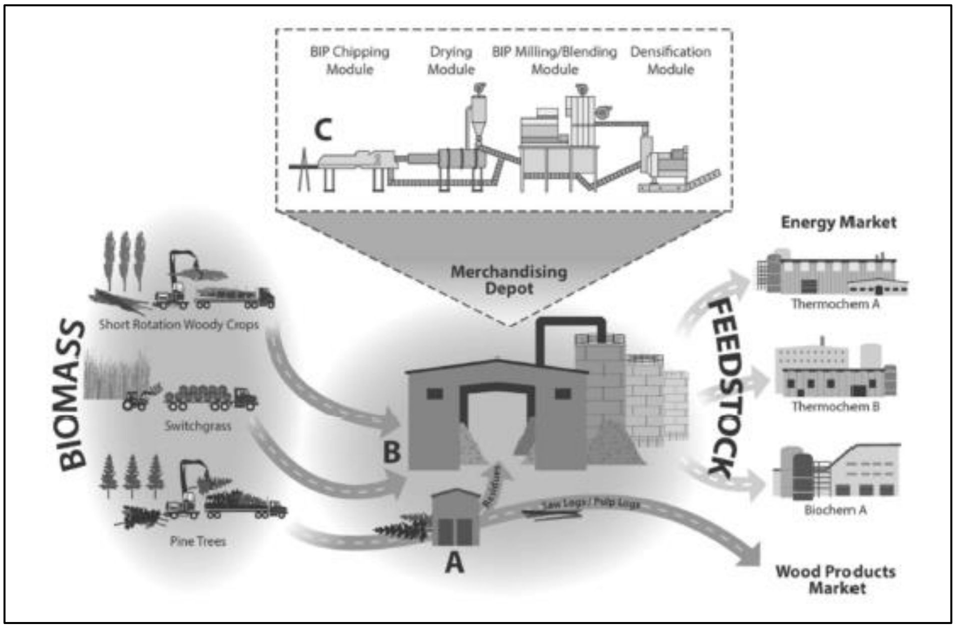


Figure 7. Bio-depot concept for standardized feedstocks (Credit: T. Rials and R. Longmire) (Platzer, 2016).

Yancey et al. (2013) showed that drying herbaceous feedstocks takes less energy than woody biomass. Dried biomass have reduced dry matter loss and degradation (Lamers et al., 2015). Biomass is either dried passively after harvest on the field or actively with additional machinery. Passive drying is a cost-efficient drying method, because additional drying equipment is not required. However, this method is limited through regional weather differences, attainable final moisture content, and occupancy of possible feedstock production areas. Studies showed that the optimal moisture content for conventional pellets for woody biomass is 5-10% and for agricultural grasses 10-20% (Stelte et al., 2012). Rotary dryers are typically used in bio-depot concepts for active drying (Tumuluru et al., 2016). This type of dryer effectively produces evenly dried particles to meet the specification limits. To ease the process of meeting the specification limits feedstocks are blended. The scope of this process is to mix more expensive feedstocks with good attributes with cheaper feedstocks with bad attributes. For example, blending forest residues (*e.g.*, pine) with an ash content of 2.6% and switchgrass with 5.8% leads to improvement of biorefinery supply through a higher quantity of less expensive feedstock types (Ray et al., 2017). The final preprocessing step is densification. Densified feedstocks are easier to handle, have a better particle size distribution and uniformity, improved compositional quality, and have properties to meet the set conversion specification limits. Densification systems such as the pellet mill, screw extruder, or piston press are commonly used to produce uniform products. The following requirements for moisture content and particle size exist for densification systems: pellet mill 10 – 15% and <3 mm, screw extruder 10 – 15% and < 20 mm, and piston press 10 – 15 % and 6 – 12 mm (Tumuluru et al., 2011).

Feedstocks

Cellulosic feedstocks, such as forest residues and Switchgrass, are major sources for cellulosic ethanol production and may be able to substitute 30% of the current petroleum-based fuel consumption (Perlack et al., 2005). This feedstock type has advantages properties for biomass to biofuel conversion. These properties are a) abundant in occurrence, b) non-edible, c) do not interfere with other market segments, and d) their chemical properties can be adjusted through blending or preprocessing (Hoekman, 2009).

Forest Residues

Wood compared to perennial grasses (*e.g.*, Switchgrass) has great properties for biofuel production such as lower ash content. For example, the ash content for pine wood is one percent compared with 5.8% for Switchgrass straw (Tao et al., 2012). Given the current poor market situation for biofuels, the biofuel production industry cannot economically compete against traditional industries, such as pulp industry or other forest product industries, that rely on roundwood (Galik et al., 2009). However, these harvest operations generate a significant number of residues which can be used for biofuel production.

Compared to normal logs, forest residues have poorer quality, smaller diameters, and are bulky. Forest residues are defined as byproducts from harvest operations such as tree tops, branches, bad quality logs, and non-merchantable stems (Moriana et al., 2015). Currently, 93 million dry tons of forest residues are removed from United States forests annually (Smith et al., 2009). This removal increases the utilization ratio of the United States forest use and increases revenue sources for the forest suppliers (IEA Bioenergy, 2007).

Forest residues are available in certain regions of the U.S. (Figure 8) (Roberts, 2014). However, the sustainability of biomass removal from forests depends on the conditions of each collection site (Nettles et al., 2015). Thirty-five percent of logging residues and 50% of other forest related removals (*e.g.*, branches, etc.) have to be left on site to maintain soil quality (Roberts, 2014). Large removals of residues from low quality sites, such as loblolly pine, can lead to less productivity in the future (Cantor and Rizy, 1991).

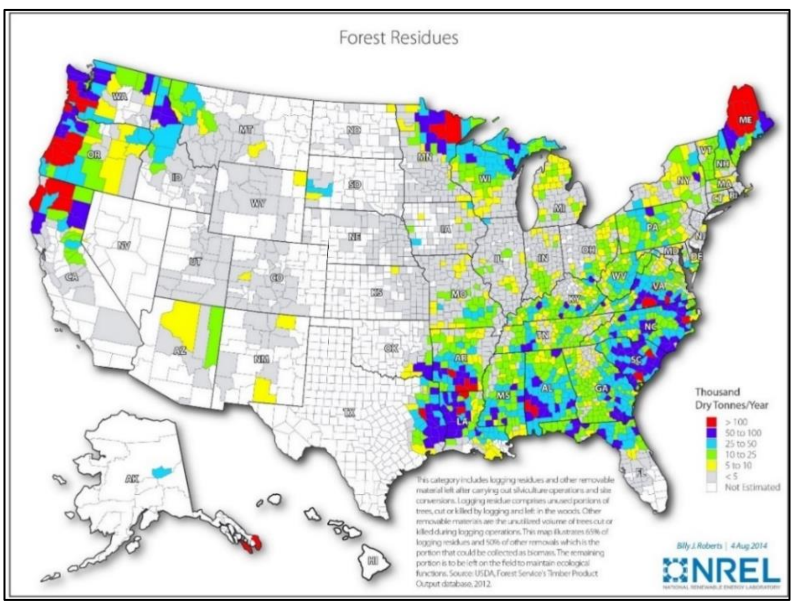


Figure 8. Forest Residues - United States of America (Roberts, 2014).

Unlike perennial grasses or agricultural residues, forest residues can be harvested annually. Forest residues are usually collected from wide areas and stored in piles at the roadside (Cambero et al., 2015). Afterwards, the biomass is dried, preprocessed in bio-depots, and transported to biorefineries (Cambero et al., 2015). However, forest residue collection should occur simultaneously with the harvest operations of roundwood to generate a more efficient and economical supply chain stream (Schnepf, 2011). The properties of freshly collected forest residues are not suitable for biomass conversion (Schnepf, 2011). Furthermore, different wood species such as pines, willows, or hybrid poplars impact biofuel conversion performance through the difference in quality characteristics, *e.g., ash content, moisture content, and particle size* (Schnepf, 2011). Studies have indicated that biofuel production from forest residues generate the best outcome using biochemical and thermochemical conversion technologies (EPA, 2007).

Mill residues like edgings, trimmings, or sawdust can also be used for biofuel conversion. However, most of the sawmill residues are already used by the mills itself for producing pellets, other wood composite products, or for energy (Douglas, 2010).

Switchgrass

Switchgrass (*Panicum virgatum* L.) is a warm-season perennial herbaceous grass species, which developed from a forage crop to an energy crop (Zegada-Lizarazu et al., 2012). Based on comparative studies, conducted by the Oak Ridge National Laboratory, Switchgrass is considered a model species for biomass energy production (Vogel et al., 2010). This status was based on features such as low establishment costs, soil conservation benefits and high adaptability to poor soil quality, wildlife enhancement, and the ability to be harvested with conventional agricultural equipment (Vogel et al., 2010, McLaughlin et al., 2002). Switchgrass occurs in all areas East of the Rocky Mountains (Figure 9) in two major ecotypes, upland and lowland switchgrass (Casler et al., 2011). The roots of both types reach a depth of 3 m (Ma et al., 2000) and a height for upland 1.5 m – 2 m and for lowland ecotypes 3 – 4 m (Moser and Vogel, 1995).

Studies have indicated that lowland Switchgrass yield up to one and a half times more biomass than upland Switchgrass (Parrish et al., 2012). Switchgrass reaches its full potential in the third year after seed establishment given the enhanced root development

(McLaughlin and Adams Kszos, 2005). Furthermore, Switchgrass can be grown on marginal croplands and on areas suitable for the Conservation Reserve Program, *i.e.*, *marginal cropland* (Vogel et al., 2010). Switchgrass harvest has higher labor costs due to seasonal availability (Bassam, 1998). Field-drying is also required to reduce moisture and reduce loss from degradation in long-term storage (Mitchell and Schmer, 2012).

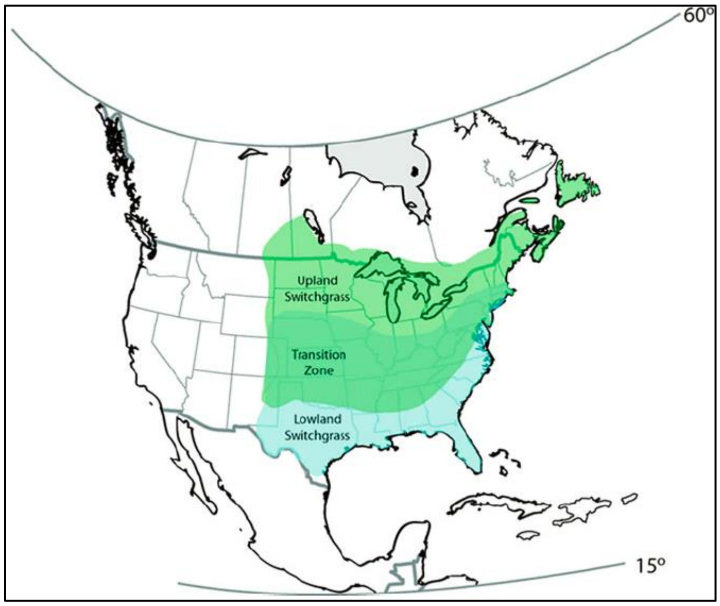


Figure 9. Native ranges of upland and lowland Switchgrass ecotypes in North America (Casler et al., 2011).

Variation of Feedstock Quality Characteristics

Variation of feedstock quality characteristics has significant impact on the performance of all units in the biomass to biofuel production (Williams et al., 2015). Historically, the biomass to biofuel supply chain is based on existent supply chain systems from different industries, such as agriculture, logging, or food production. In addition, the same biomass / feedstock specifications were inherited from those systems (Kenney et al., 2013). However, the success of biofuel production is based on a steady, cost efficient, and controlled quality feedstock supply (Thompson et al., 2014). In recent decades the majority of research tried to optimize and reduce costs of biomass logistics in the supply chain through machine development and material loss reduction across the supply chain (Zandi

Atashbar et al., 2017). But there is an absence in the literature on focusing on the importance of variation of feedstock quality characteristics.

Feedstock quality characteristics can be categorized as follows: physical on a macroscale, structural on a microscale, and compositional on a molecular scale (Li et al., 2016). First, physical characteristics such as feedstock type, particle size and shape, or moisture content impact feedstock processing and handling. Second, structural characteristics such as cellulose crystallinity, affect selection of conversion technology. Finally, compositional characteristics such as ash content impact feedstock selection and production yield (Li et al., 2016). Due to the impact on the performance of biofuel production from biomass, ash content, moisture content, and particle size are set as the key quality characteristics for the simulations in this research.

Ash Content

Ash content has a negative impact on the biomass to biofuel conversion performance (US Department of Energy, 2014). Ash in biomass feedstock originates from either the natural physiology of the plant or through contamination with soil or rocks, *e.g., forest residues versus roundwood* (Lacey et al., 2016). Natural ash in plants is either associated with structural ash in cell walls or vascular in cell extracts (Kenney et al., 2013). In addition, studies showed that the mechanical processing at the harvesting operations introduce ash content into the biomass. Ash content varies between and within biomass types (Table 2), *e.g., woody biomass compared to herbaceous plants and roundwood compared to woody residues* (Tao et al., 2012).

Ash in any form within feedstocks has negative impact on biochemical and thermochemical conversion technologies. Studies have indicated that corn stover has a higher ash content which reduces the effectiveness of pretreatment processes and displaces carbohydrate, which is crucial for the biofuel conversion (Weiss et al., 2010). According to Kenney et al. (2013) there is no specification limit for ash content using the biochemical conversion process. In contrast, for pyrolysis-based thermochemical conversion processes the specification limit is one percent (US Department of Energy, 2014). Biomass with high ash content negatively impacts the pyrolysis process through the creation of slag formation within the combustion process and decreased efficacy of the catalysts used (Kenney et al., 2013). Preprocessing, such as fractionation or the use of specific harvest equipment, can reduce the ash content in the biomass (Shinners et al., 2012) and therefore increase the biofuel yield. Furthermore, Shinners et al. (2012) illustrated that biomass harvested with multi-pass equipment has a higher ash content from increased soil contact relative to biomass collected with single-pass equipment.

Table 2. Mean values and ranges for ash content of selected lignocellulosic biomass feedstocks.

Feedstock	Mean ash (%)[1]	Reported range (%)
Herbaceous		
Switchgrass straw	5.8 (21)	2.7 – 10.6
Woody		
Pine wood	1.0 (40)	0.1 – 6.0
Pine residue	2.6 (4)	0.3 – 6.0
Spruce wood	0.8 (5)	0.3 – 1.5
Spruce residue	4.3 (2)	2.2 – 6.4
Willow wood	1.5 (18)	1.0 – 2.3
Willow residue	2.0 (1)	2.0 – 2.0

[1] Sample number of mean values in parenthesis
Data taken from (Tao et al., 2012); inspired by (Kenney et al., 2013)

Moisture Content

Biomass moisture is a crucial cost driver for biofuel production. Excessive moisture negatively affects storing, transporting, handling, and feeding. Biomass handling and feeding becomes tedious with increased moisture content, because the cohesive strength of the material increases and therefore can plug feeders (Dai et al., 2012). Emery and Mosier (2012) showed that dry matter loss for aerobic stored biomass increases with moisture content. Furthermore, wet biomass decreases truck utilization for transportation, *i.e., transportation of less biomass and more water* (Eggink et al., 2018). Biomass moisture affects not only biofuel conversion performance, it also affects grinding energy and execution (Tumuluru et al., 2014) which indirectly impacts the conversion performance (Williams et al., 2015). The specification limits for the moisture content for woody residues (Keefe et al., 2014) and herbaceous biomass depend on the final conversion technology. Tumuluru et al. (2011) summarized optimal moisture content specifications limits for different densification equipment's, *e.g., pellet mill with 10-15% or a piston press with 10-15%, etc.*. Technical targets for main supply chain operations were introduced by Jacobson et al. (2014) (Table 3).

Table 3. Technical targets for typical supply chain operations for woody residues and Switchgrass.

Supply Chain Operation	Woody Residues	Switchgrass
Harvest and Collection	40%	5-10%
Field Storage	30%	20%
Transport	30%	20%
Drying	30%	30%
Densification	19%	19%
Blending of Pellets[1]	9%	9%

[1] Feedstocks were individually pelletized and blended based on final blend-formulation
Targets were taken from the Idaho National Laboratory "Feedstock Supply System Design and Analysis" – Case study for thermochemical conversion - (Jacobson et al., 2014)

Particle Size

Particle size defines the flowability and bulk density of cellulosic feedstocks which crucially impact the efficiency of the biomass supply chain and the attainable biofuel yield through biomass to biofuel conversion processes (Bitra et al., 2009, Miao et al., 2011). Comminution – particle size reduction – is vital to increase flowability and bulk density of cellulosic raw material to increase supply and conversion process efficiency (Hess et al., 2009, Miao et al., 2011). The location of particle size reduction determines the success of the whole supply chain; particle size of cellulosic biomass is best modified at an early stage (Meunier-Goddik et al., 1999). Furthermore, technology, logistics, and economic feasibility of the comminution operation are determined by the supply chain design (*e.g.,* type of storage or transportation, etc.) and conversion technology used in the biorefinery

(Lam et al., 2008). The associated high energy consumption and processing costs (*e.g.*, required pre-drying due to high moisture in biomass) are problematic for particle size reduction (Schell and Harwood, 1994), while generating low-value products (Himmel et al., 1985). Comminution of biomass is generally required for all conversion technologies (Williams et al., 2015). Biochemical conversion process is more tolerant of particle size variation than thermochemical conversion processes (Kenney et al., 2013). However, neither fines nor over-sized particles are desirable for an optimal execution of the conversion process (Kenney et al., 2013). Particle size and distribution depend on the milling equipment used, typically either hammer mills or knife ring flakers are used. Particles produced from hammer mills tend to be finer than from knife ring flakers for the same screen size (Kenney et al., 2013). Specifications and targets of particle size reduction are set by the requirements of the end-users (Igathinathane et al., 2008). Furthermore, some studies suggest that particle size has no influence (*i.e.*, no significant correlation) on the sugar production from cellulose (Vidal et al., 2011), others showed that reduced biomass particles have greater digestibility than bales for the conversion process (Hess et al., 2009).

Continuous Improvement

'Kaizen' is a popular Japanese term that is defined as small steps toward continuous improvement (CI). Kaizen is a company-wide philosophy which utilizes many tools to enhance the performance of the enterprise (Singh and Singh, 2015). CI was a philosophy developed by Deming (1982, 1986, 1993) and is defined as a "*never-ending process to improve the current state of the worker, process, production, or enterprise".* Juran (1989) redefined CI as '*Total Quality Management*' (TQM) which describes incremental improvement through participation of all entities and people of an organization (Bhuiyan and Baghel, 2005). The goal of any improvement philosophy is to drive defects towards zero by reducing variation around the target value (Chen, 2004). The *'Toyota Production System'* or TPS (Ohno, 1988) which was redefined by Womack (1996) as *'Lean Thinking'* focuses on the elimination of waste in an organization, *e.g., excessive variation is defined as waste in TPS or Lean Thinking.* Six-Sigma quality (Harry and Schroeder, 2000)

encompasses all the previously defined improvement philosophies and also focuses on a methodical approach to using statistical methods to improve organizations and improve quality. The name *'Six-Sigma'* is defined in this philosophy as having a natural tolerance that is six standard deviations within specifications, or only producing 3.4 out of one million parts that are defective. George (2002) combined the TPS (or Lean) and Six-Sigma philosophies and further redefined continuous improvement as *'Lean Six-Sigma'* or LSS. The core method in all of the aforementioned improvement philosophies is the use of statistical methods to quantify variation and identify sources of variation influencing variation in processes; with the ultimate goal of variation reduction, process/product improvement, and lower costs (Taguchi et al., 1988).

The Protagonists of the 20^th Century Quality Revolution

The quality revolution of the 20th century began with the invention of the control chart by Dr. Walter A. Shewhart. Shewhart's breakthrough philosophy was that quality control can only be ensured by eliminating process variation (*i.e.*, prevention) and not just by inspection only and removing defective products from finished batches (Shewhart, 1931, Shewhart, 1939). After introducing his ideas at the Bell Telephone Laboratories Dr. Shewhart hired an inquisitive and ambitious Ph.D. student called W. Edwards Deming in 1927. Deming, fascinated by Dr. Shewhart's thinking, saw the potential of Shewhart's ideas on statistical methodologies to improve manufacturing and applied them in a greater management context (Tsutsui, 1996). Deming's 14 points for management, the seven deadly diseases of management, and the Shewhart Cycle are critical contributions of Dr. Deming for quality control in the 20th century to increase the performance of Japanese and U.S. manufacturers (Deming, 1986, Deming, 1993).

Unfortunately, the potential of Dr. Deming's ideas and views on quality control were unrecognized by American managers after the end of the second world war (Tsutsui, 1996). Unrecognized by U.S. companies, Drs. Deming and Joseph M. Juran were invited by the Japanese Union of Scientists and Engineers (JUSE) to give lectures about their teachings to help the emerging Japanese automobile industry gain competitiveness in world markets. Dr. Juran like Deming, suggested that only management can improve the state of the production (Juran and Gryna, 1951, Juran and Gryna, 1993). Meanwhile, Japanese engineers and JUSE members such as Taiichi Ohno (*i.e.*, TPS) and Genichi Taguchi (*i.e.*, Taguchi loss functions and robust product design) developed methods to continuously improve production by quantifying and reducing variation (Ohno, 1988, Taguchi, 1993, Taguchi et al., 2004). After the U.S. automotive industry lost significant market share to Japanese auto manufacturers in the 1970s, Dr. Deming appeared on an NBC documentary titled, *"If Japan can… Why can't we?".* Many believe the June 24th, 1980 NBC broadcast was the genesis for the "American Quality Revolution".

Key Methods in Continuous Improvement

Statistical Process Control

The first phase in continuous improvement is defining the state of the process. Control charts are considered the key statistical method for SPC and continuous improvement (Deming, 1986, Grant et al., 1994). Control charts are fundamental to SPC in that the stability of the process is quantified and is visualized. The invention of control charts by Walter Shewhart in the 1920s and applied at Bell Laboratories in the 1930s was the genesis for the development of SPC (Shewhart, 1931, Wheeler and Chambers, 1992). SPC uses the control charts (a statistical '*prediction interval*') to visualize variation and predict of future process outcomes[1]. The control chart quantifies and distinguishes variation as two-types: 1) common-cause variation; and 2) special-cause variation or '*events*' (Figure 10). Monitoring variation using the control chart can prevent the manufacture of defective product known as '*scrap*' and reduce rework (Young and Winistofer, 1999).

---

[1] It is important to distinguish between a statistical 'prediction interval' and 'confidence interval'. The control chart is an analytical technique for prediction on data that is continually changing and is defined by the control limits which are: $\bar{X} \pm 3 \times s$. The confidence interval is an enumerative technique on a reference frame or sample that does not change, and are typically defined assuming unknown variance by: $\bar{X} \pm t_{\frac{a}{n},n-1} \times \left(\frac{s}{\sqrt{n}}\right)$. Prediction intervals are typically wider than confidence intervals.

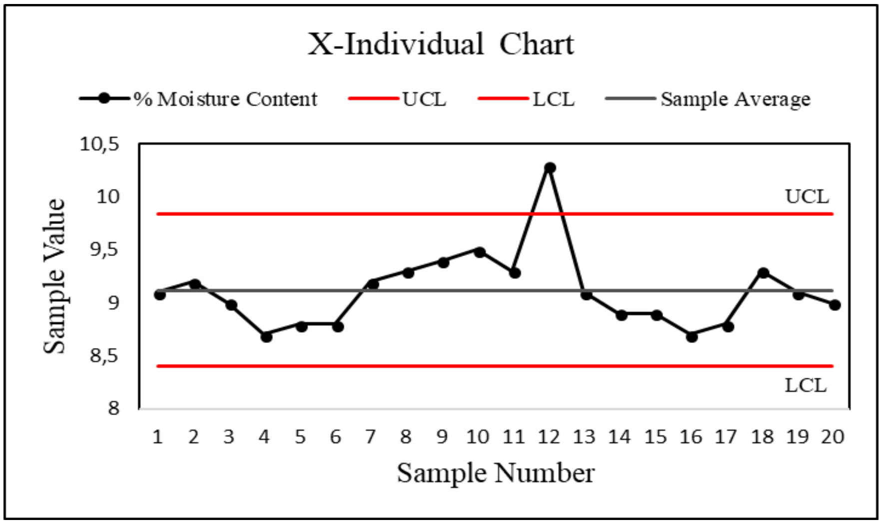


Figure 10. Example control chart: X-individual chart with one outlier.

Common-cause variation is natural variation in a process, product, or material and is a stable, consistent pattern that leads to prediction of the process. Special-cause variation is unstable variation created by an event, *e.g., shift change, raw material change, etc.*

Optimization of a production process can only take place when the process is stable and does not suffer from special cause variation, it is crucial to eliminate special-cause variation first. Control limits are not specifications limits (engineering tolerance) and are approximately ± three sigma (σ) from the process average ($\bar{x}$) (Young and Winistofer, 1999). Control limits contain approximately 99.7% of the variation and assume a normal or Gaussian distribution of the data (Sauers, 1999). There are many different types of control charts, and depending on the application and sampling, the upper and lower control limits (UCL / LCL) are computed using different equations (Wheeler and Chambers, 1992).

Originally, four control run rules were introduced by the Western Electric Company (1956) and later updated to eight by Lloyd S. Nelson (1984) to detect special-cause variation in control charts, which are the following:

1. One point is more than three standard deviations from the mean, *i.e., outlier indicates a process out of control.*
2. Nine (or more) points in a row are on the same side of the mean, *i.e., indicates a shift in the mean.*

3. Six (or more) points in a row are continually increasing (or decreasing), *i.e., indicates a trend.*
4. Fourteen (or more) points in a row alternate in direction, increasing then decreasing, *i.e., indicate at least two different data sets.*
5. Two (or three) out of three points in a row are more than two standard deviations from the mean in the same, *i.e., indicates a shift in the mean.*
6. Four (of five) out of five points in a row are more than one standard deviation from the mean in the same direction, *i.e., indicates a shift in the mean.*
7. Fifteen points in a row are all within one standard deviation of the mean on either side of the mean, *i.e., a higher variation would be expected.*
8. Eight points in a row exist, but none within one standard deviation of the mean, and the points are in both directions from the mean, *i.e., indicate at least two different data sets.*

Shewhart distinguished between control charts for measurement data and attribute data. Measurement data come from continuous measurements and are considered a real number, *e.g., heights, densities, moisture content, physical dimensions, etc*. (Table 4). Attribute data are integers and are data, such as number of rejects, blemishes, etc.

The previous review of literature related to control charting is meant to be an overview for the practitioner and sets the stage for a fundamental method of this thesis.

Toyota Production System or Lean

Lean manufacturing describes tools and principles for systematic and continuous improvement of manufacturing and service processes by eliminating waste with the goal to elevate the enterprises success. Lean manufacturing, originally termed "*The Toyota Production Systems*" (TPS), was invented in the 1950s by Taiichi Ohno of the Japanese automobile company Toyota Motor Corporation and was designed to overcome limitations in competing with U.S. automobile enterprises (Ohno, 1988, Sundar et al., 2014). TPS or Lean focuses on the elimination of waste *"Muda"*, variation *"Mura"*, and over-burdening of systems and workers *"Muri"* (Radnor and Leseure, 2010). Lean defines seven major types of waste which are overproduction, waiting, transport, inappropriate processing, unnecessary inventory, unnecessary motion, and defects (Wilson, 2010). These types of wastes should be understood in terms of value-added and non-value-added activities (wastes) to the final product based on the customers view. Value-added activities help converting raw-material or semi-finished products to its finished state and are actions the customer wants to pay for, while non-value-added activities are wastes and unnecessary actions in the conversion process of a product (Hines and Rich, 1997). The key metric for improvement in Lean is '*Value Stream Mapping*' which relies on the '*Value-Added Ration (VAR)*' to determine if a process has been improved. Value stream maps highlight the process as a flow chart, define processing time into either 'value-added' or 'non-value-added' times, *e.g., cycle time, change over times, etc.* (Rother and Shook, 1999).

$$VAR = \frac{Time\ used\ for\ the\ Process}{Total\ Process\ Cycle\ Time} \qquad [1]$$

Flow Charts

Flow charts are a useful tool during the initial root-cause analyses phase of continuous improvement. They are helpful to visually describe a process or production. A process flowchart shows the logical sequence of activities executed to produce a product. The great advantage of flowcharting is the quick identification of process steps which should be eliminated (Srinivasan, 2011). Streamlining a process is only possible through the identification, elimination, or at least reduction of non-value-added activities. Usually standardized symbols are used to represent certain type of actions (Figure 11).

Pareto Charts

Pareto charts are a method for visualizing defects or assignable events occurring in the process (Juran and Gryna, 1951). Most successful continuous improvement efforts use the Pareto Chart to identify the critical variable inducing variation in the process. Adapted from the "80/20-rule" invented by the Italian economist Vilfredo Pareto 80% of

Table 4. Common univariate control charts for measurement and attribute data (Wheeler and Chambers, 1992, Young and Winistofer, 1999).

Control Chart Type		**Central Line**	**Control Limits**	**Purpose and when to use**
Measurement Data				
Subgroup n = 1	X-Individual	$CL_X = \bar{X}$	$UCL_X = \bar{X} + 2.660\,\overline{mR}$ $LCL_X = \bar{X} - 2.660\,\overline{mR}$	Assessment of long-and short-term process variation – periodically collected data (organization of data in rational manner)
	Moving Range	$CL_R = \overline{mR}$	$UCL_R = 3.268\,\overline{mR}$	Assessment of stability of short-term process variation – slowly changing process
Subgroup n > 1	X-bar	$CL_{\bar{X}} = \bar{\bar{X}}$	$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\,\bar{R}$ $LCL_{\bar{X}} = \bar{\bar{X}} - A_2\,\bar{R}$	Assessment of stability of the location of the process relative to its target – historical summary and organization of data into rational subgroups
	Range	$CL_R = \bar{R}$	$UCL_R = D_4\bar{R}$ $LCL_R = D_3\bar{R}$	Assessment of stability of the process variation within and between subgroups – historical summary and organization of data into rational subgroups
Attribute Data				
Binomial data	np chart	$CL_{np} = n\bar{p}$	UCL_{np} $= n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$ LCL_{np} $= n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$	n constant – all samples have the same sized areas of opportunity – counts bad and good samples
	p chart	$CL_p = \bar{p}$	$UCL_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$ $LCL_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$	n variable – Areas of opportunity changes from sample to sample – counts bad and good samples
Poisson data	c chart	$CL_c = \bar{c}$	$UCL_c = \bar{c} + 3\sqrt{\bar{c}}$ $LCL_c = \bar{c} - 3\sqrt{\bar{c}}$	a constant – all samples have the same sized areas of opportunity – used to count bad samples in complex products
	u chart	$CL_u = \bar{u}$	$UCL_u = \bar{u} + 3\sqrt{\frac{\bar{u}}{a_i}}$ $LCL_u = \bar{u} - 3\sqrt{\frac{\bar{u}}{a_i}}$	a variable – Areas of opportunity changes from sample to sample – used to count bad samples in complex products

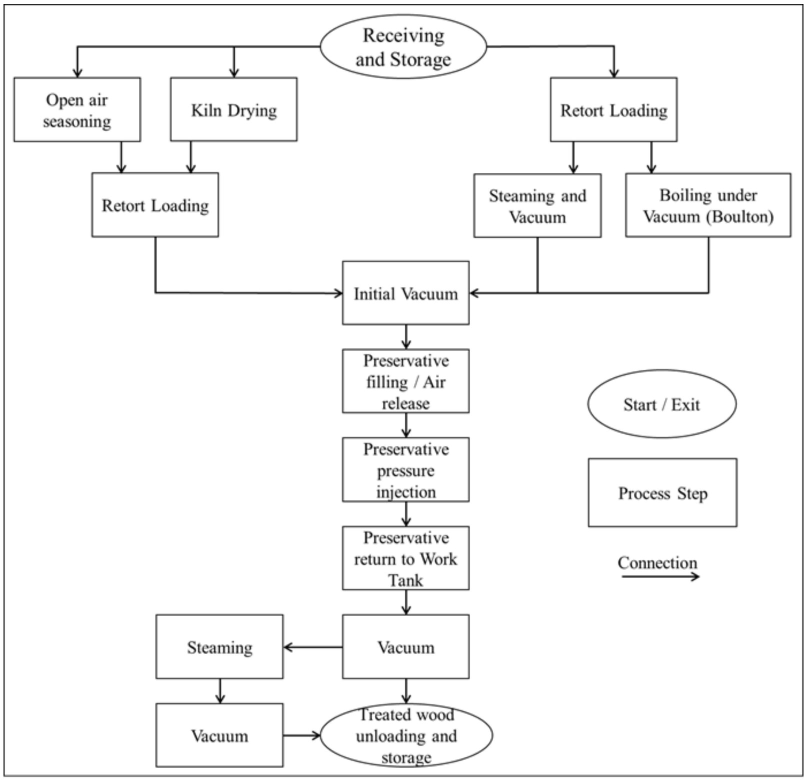


Figure 11. Example flow chart: Full-cell pressure treating process for treated lumber (Institute, 1999).

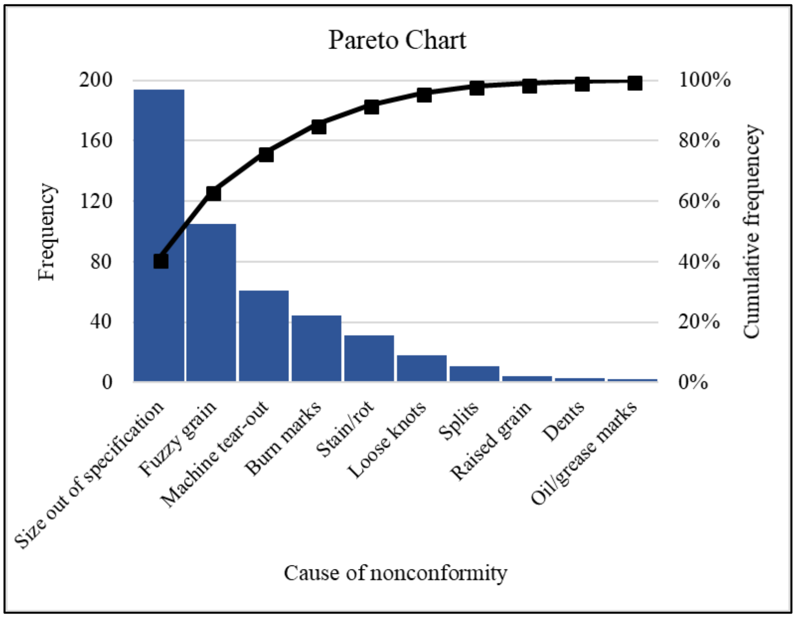


Figure 12. Pareto Chart for causes of nonconformity of a wood product (Leavengood and Reeb, 2002).

the variation in a process originates from 20% of the causes (Wilkinson, 2006). The Pareto chart is a histogram (Figure 12) where causes or defects are organized by largest frequency from left to right. The identified causes for the problem are represented by bars on the horizontal axis; the cumulative contribution by the causes are represented on the vertical axis via a line. This technique easily identifies the main cause of the problem.

Ishikawa Diagrams

Once the main problem has been identified on the Pareto Chart, the typical next step is to develop Ishikawa or "cause-and-effect" or "fishbone" diagrams (Figure 13). Ishikawa used the diagram in organized brainstorming sessions with workers in the automobile industry in Japan to list all possible causes influencing the variable being study (Ishikawa, 1986). As many authors have noted, identification of potential sources should be done in group work of production workers and engineers to gain optimal result (Doshi et al., 2012). Usually, these sources are grouped in the following five different categories: methods, machines, people, materials, and environment (Doshi et al., 2012).

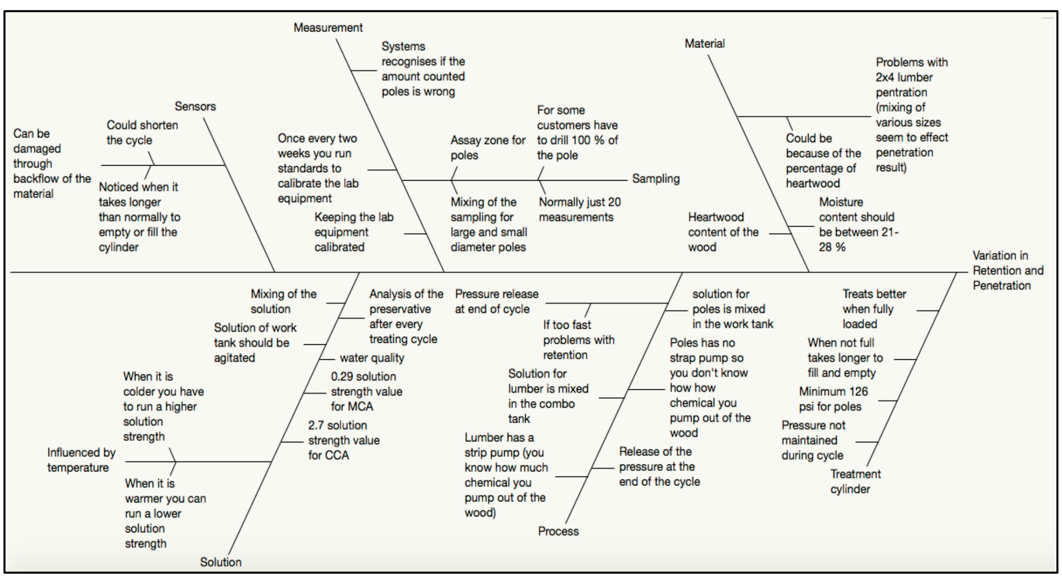


Figure 13. Example Ishikawa diagram for "Variation of Retention and Penetration of Treated Wood" (Hamernik, 2018).

Cause Mapping

Cause mapping is an expansion of the Ishikawa diagram for cause-and-effect analysis (Scavarda et al., 2004) and investigates problems closely linked to the organizations main goals. The key premise for cause mapping is system thinking, *i.e., every system has parts which are connected and interact with each other* (Zhu, 2008). Furthermore, unlike the Ishikawa diagram a cause map focuses on the cause-and-effect relationship and not specific categories, *i.e., each effect has a cause and each cause has an effect* (York et al., 2014). The cause map starts on the left with the defined problem placed in so-called effect boxes (Figure 14). The question *'Why?'* is asked to identify the cause supported by clear evidence of the effect. This scheme is repeated for each effect to create a detailed cause map of the system. Cause mapping allows for a more specific and detailed cause-and-effect analysis than the Ishikawa diagram. Additional to the Ishikawa-diagram and cause-mapping asking the question why five times is another root-cause-analysis tool from Lean. The Five-Why technique helps to find the root cause, not symptoms, of the problem and identifies their relationships.

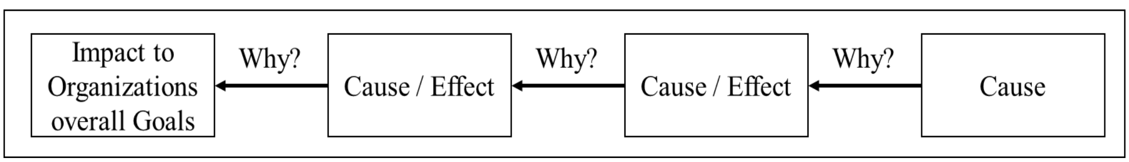


Figure 14. Typical scheme for cause mapping; cause-and-effect analysis.

Summary of Lean or TPS Methods

This section is intended to give a brief summary of important Lean (TPS) methodologies. These methods aid the continuous improvement process through organizing the workspace, streamlining the production flow, and reducing non-value-adding activities, *i.e., elimination of waste*. More information about these tools can be found in the cited literature (Ohno, 1988, Srinivasan, 2011, Wilson, 2010).

The **5S-Methodology** is used to systematically improve the workplace by removing unnecessary equipment and increasing organization through visual aids. Each step of the methodology is expressed as a Japanese term starting with S (Al-Aomar, 2011): *Seiri* ("sort" by removing unnecessary equipment and material), *Seiton* ("set in order" by organizing the workplace with visual aids), *Seiso* ("shine" by cleaning the workplace), *Seiketsu* ("standardize" by documenting working methods or using standardized procedures / equipment), and *Shitsuke* ("sustain" by continuously applying this technique).

Standardized work allows the application of best practices in the workplace. Standardizing procedures improves consistency of process execution by reducing variation (Emiliani, 2008). **Mistake proofing** (*Poka Yoke*) equipment and processes further increases product quality by integrating mechanics and sensors to immediately detect errors.

Lean methods such as *Just-in-Time*, *Jidoka*, *Heijunka*, and *Kanban* aim to streamline the production by creating a smooth flow of the produced items. **Continuous flow** is a manufacturing where the materials or products run through the production without or only minimal buffers. Reduced lead times, inventory, and smaller changeover times are associated with continuous flow.

Heijunka, mixed-model scheduling, is used to distribute production capacity equally on each product by reducing batch size. Smaller batch sizes lead to smaller lead times which allows the production to better meet customer demand. However, smaller batch sizes lead to more necessary changeover setups. **SMED ("**Single Minute Exchange of Dies") was developed, by Shigeo Shingo in the 1950s, to exactly handle the increasingly smaller becoming production lot sizes (Ulutas, 2011). SMED methodology aims to reduce changeover time to less than 10 minutes by applying the following three main steps: 1) execute all setup up steps externally if possible; 2) convert internal setup to external setup; 3) streamline the changeover, *i.e., standardize all required procedures for the changeover* (Ulutas, 2011).

Jidoka, "*autonomation with a human touch*" (Ohno, 1988), is the first key element for the success of TPS. Jidoka describes the partial automation of production systems combined with defect detection systems. This method allows workers to monitor several processes at the same time and detect quality issues immediately.

Takt-time, stems from the German word Takt (rhythm), is a means to pace the production of each item. Takt-time is simply a ratio of the available time per period and product demand per period, *i.e., allows to compare actual production with the target of the product.*

Kanban, a key technique of lean for continuous improvement, regulates the continuous flow through emphasize on the pull replenishment principle, *i.e., a product should only be produced if customer demand exists.* Signal cards are used to indicate the need of products or materials. Kanban reduces inventory and prevents overproduction.

Finally, all those aforementioned techniques enable **Just-in-Time** (JIT) for generating continuous flow. JIT strongly emphasizes the pull principle introduced with Kanban. Parts should only be produced with raw materials arriving at the right time with the right amount at the right place for the right product. Taichii Ohno mentioned JIT is the second key element for the success of TPS. The advantages are reduced inventory and space requirements.

Theory of Constraints

Eliyahu M. Goldratt developed Theory of Constraints (TOC) to provide a thinking concept on how to tackle material or managerial limitations in manufacturing to greatly improve the systems performance (Srinivasan, 2011). These production limitations, bottlenecks, essentially constrain the process execution and as a result restrain the overall success of the enterprise (Blackstone, 2001). For example, the constraint for the Switchgrass supply chain is the harvest and collection operation due to seasonal availability of the biomass. A perfect enterprise would have no constraint and would make infinite profit (Blackstone, 2001). Therefore, in TOC the success of an organization is based on how well all processes work together. This theory provides a five-step approach to solve the constraints individually and implements an additional way for continuous improvement of a system (Goldratt, 1990, Rand, 2000, Srinivasan, 2011).

At first, the manager or engineer should start with (1) *identifying the system's constraint(s)*. The choice of constraint should be based on the constraints impact on the performance of the production. Constraints can be either physical, for example limited machine capacity or material variation or based on policy. Policy constraints can either be created from poor process methodology or by flawed design of regulations and rules in an organization. After the constraints identification there should be a discussion on (2) *how to exploit the system's constraint(s)*. Physical constraints should be used as effectively as possible. In contrast, a flawed policy should be eliminated and replaced with an improved new policy. (3) S*ubordinate everything else to the above decision* for achieving maximum success with the current production environment. By subordinating all resources to the main constraints needs allows to maximize its output and essentially improve the total systems performance. This is possible since non-constraint resources have productive and non-productive capacities; optimal used non-constraint resources have no impact on the performance. If the identified (1) and exploited (2) (3) constraints are still existent it is crucial to (4) *elevate the system's constraint(s)* to generate more company profit. Elevating means to find actions to improve the systems overall performance. For example, if resource (machine) capacity is limiting the production output buying another machine to gain increased production capacity would elevate the system. Thus, another constraint in the production will arise and will form the new constraint - (5) *if a constraint was broken in a previous step, go back to step 1*. Step 5 implies that TOC should be seen and executed as a continuous improvement process; inertia should not allow to restrict the performance of the enterprise.

Taguchi's Quality Loss Functions

Quality loss functions are used to quantify the loss caused by variation in quality characteristics (Taguchi et al., 2004). Genichi Taguchi developed his quality loss functions to support the quality revolution for the Japanese industry (Lofthouse, 1999). The goal of quality loss functions is to quantify the loss caused by variation of product quality characteristics, such as ash content, moisture content, or particle size, in cellulosic biomass. Quality characteristics are performance characteristics, which affect the final quality of a product (Antony, 1997). In Taguchi's philosophy a production experiences loss in revenue when the product defining quality characteristic deviates from the target (Teeravaraprug, 2008). For example, if the product meets the target, the loss is zero. However, if the deviation from the target is double the experienced loss quadruples (Kim and Liao, 1994). Crucial for Taguchi's philosophy is that the financial loss will be experienced after the shipment of the product, *i.e., customer dissatisfaction through possible product repair or replenishment, which may cause reputational damage and lead to loss in market shares* (Taguchi et al., 2004).

Genichi Taguchi developed three quality loss functions: *nominal-the-best*, *smaller-the-better,* and *larger-the-better* (Teeravaraprug, 2008). In addition, the loss can either be computed for just one sample or for a set of samples. The two-sided loss function *nominal-the-best* (Figure 15) is used for quality characteristics with a known target, upper specification limit (USL) and lower specification limit (LSL), *e.g., moisture content or particle size.* The symmetrical two-sided loss function for one unit is determined as (Taguchi et al., 2004)

$$L = k \times (y - m)^2, \qquad [2]$$

while the loss function for more than one unit is

$$L = k \times [\sigma^2 + (\bar{y} - m)^2]. \qquad [3]$$

Where:

L = loss in dollars with the average $\bar{y}$ of the quality characteristic,

$\bar{y}$ = the average of the quality characteristic y, *e.g., moisture content particle size, etc.*,

m = target of the quality characteristic y,

k = proportionality constant,

σ^2 = the variance around the average $\bar{y}$.

The proportionality constant or cost constant *k* is defined as:

$$k = \frac{A_0}{\Delta_0^2} \qquad [4]$$

Where

A_0 = consumer loss at consumer tolerance,

Δ_0 = consumer tolerance.

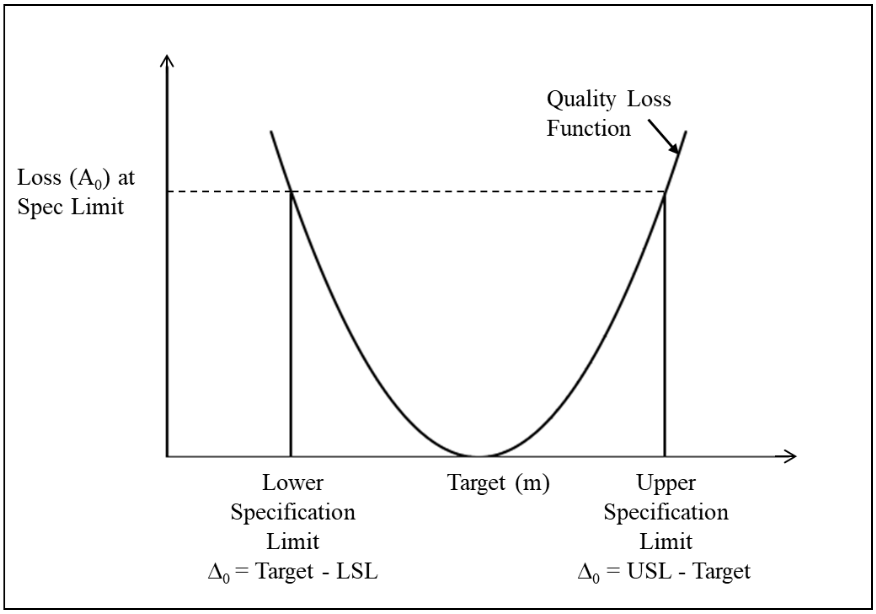


Figure 15. Taguchi's quality loss function: Symmetric *nominal-the-best*.

Equations [2] and [3] are suitable to compute the loss for symmetric specification limits, but not for asymmetric specification limits. Asymmetric specification settings exist if either the consumer tolerance (Δ_0) for USL and LSL or the consumer loss at consumer tolerance limits (A_0) are different. Kim and Liao (1994) and Liao (2010) suggested to adjust the cost constant k of Taguchi's equation [2] to represent the asymmetric specifications. Thus, the losses for values smaller than the target is computed as

$$L(y) = k_{LSL} \times (y-m)^2 \, for \; y < m, \qquad [5]$$

and for values greater than the target is computed as

$$L(y) = k_{USL} \times (y-m)^2 \, for \; y > m. \qquad [6]$$

However, equations [5] and [6] do not give sufficient information about the influence of the variation of a quality distribution. Li (2002) provides an overview of complex linear and quadratic models for the application of asymmetrical quality loss functions. To reduce complexity, this thesis will analyze if the same procedure for equations [5] and [6] can be applied to equation [3].

In contrast, the smaller-the-better loss function (Figure 16) is used for quality characteristics where minimizing the result is wanted; ideally zero. For example, ash content negatively impacts the biomass to biofuel conversion performance. The equation for the *smaller-the-better* loss function for one unit is defined as the following (Taguchi et al., 2004):

$$L = k \times y^2 \qquad [7]$$

The equation for the *smaller-the-better* for more than one unit is defined as:

$$L = k \times [\sigma^2 + \bar{y}^2] \qquad [8]$$

Where:

L = loss in dollars with the average $\bar{y}$ of the quality characteristic,

$\bar{y}$ = the average of the quality characteristic y, *e.g.*, *ash content*, *etc.*,

k = proportionality constant,

σ^2 = the variance around the average $\bar{y}$.

Where the proportionality constant k is equal to:

$$k = \frac{A_0}{y_0^2} \qquad [9]$$

A_0 = consumer loss at consumer tolerance,

y_0 = consumer tolerance.

On the contrary the larger-the-better loss function (Figure 17) is used for quality characteristics where maximizing is desired. For example, increased sugars in biomass improve the biofuel yield. The following equation defines the loss for the larger-the-better loss function (Taguchi et al., 2004).

$$L = k \times \frac{1}{n} \sum_{i=1}^{n} \frac{1}{y_i^2} = k \times \frac{1}{n} \left(\frac{1}{y_1^2} + \frac{1}{y_2^2} + \cdots + \frac{1}{y_n^2} \right) \qquad [10]$$

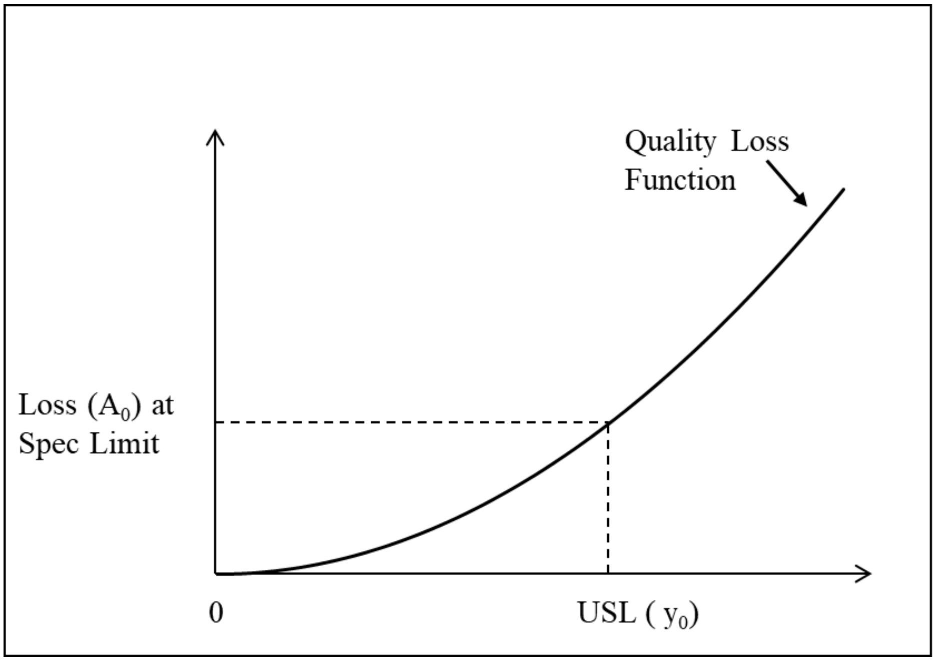


Figure 16. Taguchi's quality loss function: *Smaller-the-better*.

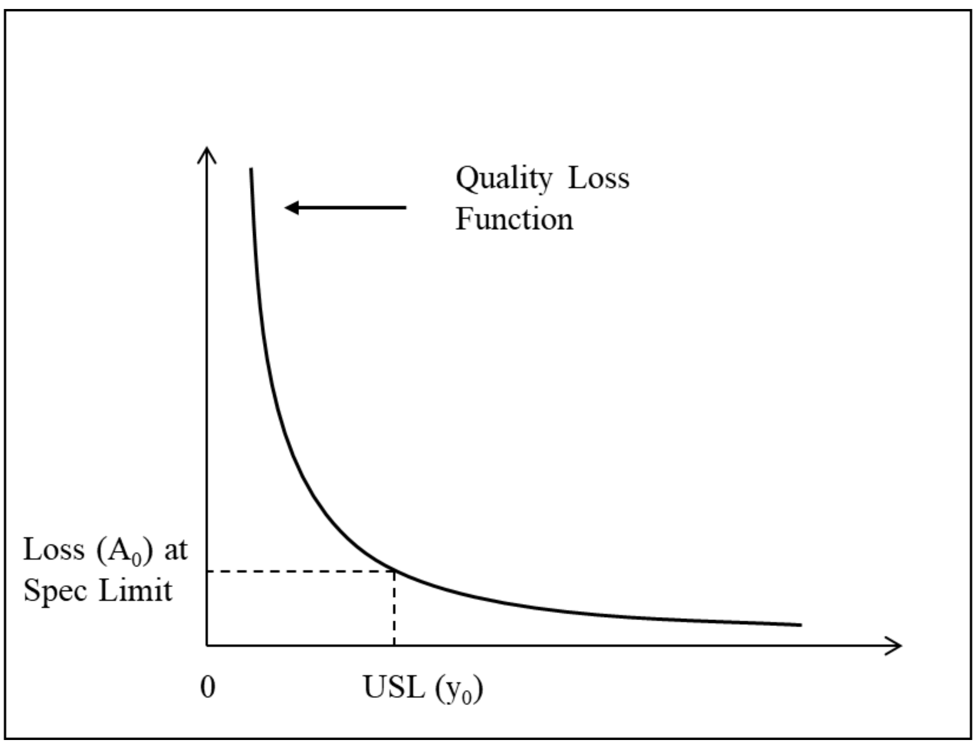


Figure 17. Taguchi's quality loss function: *Larger-the-better*.

Where:

L = loss in dollars with the average $\bar{y}$ of the quality characteristic,

y_i = ith value for the quality characteristic of y, *e.g., formaldehyde (CH_2O) emission, etc.*,

k = proportionality constant.

Where the proportionality constant k is equal to

$$k = A_0 y_0^2 \qquad [11]$$

A_0 = consumer loss at consumer tolerance,

y_0 = consumer tolerance.

The above described loss functions indicate that variation in quality in context of Taguchi's philosophy should be seen more carefully. Compared to classical quality thinking, where all products within specification limits are treated as equally good, Taguchi implies that the experienced loss is greater for higher deviations of quality characteristics (Liao, 2010).

Components of Variance

Galton's early writings on the idea of statistical studies established the framework for the concept of '*components of variance'* (Stigler, 2010). Galton's theory was that in any system variance may accumulate through the system, so that the total variance is the sum of the components. The concept of components of variance is the basis for the quantifying of the variability on the supply chain for biomass developed in this thesis.

Variance is accumulated in the following biomass supply chain example. In the case of a series system (*e.g.*, biomass supply chain) the variance of a certain quality characteristic (*e.g.*, moisture content) may have an impact on the variance of the feedstock of the subsequent steps. For example, increased moisture content of harvested biomass can have an impact on the dry matter loss. Depending on the storage type, additional moisture can be introduced by environmental influences, which increases the overall variance of the system.

Therefore, mathematically the sum of variances is defined for any series or parallel system (Montgomery, 2012). Under the assumption that the variables X and Y are random in a parallel system both variables (components) are independent. Therefore, the equation is

$$Var(X + Y) = Var(X) + Var(Y). \quad [12]$$

As mentioned earlier in a series system the variables (components) are dependent have a positive or negative influence on each other. Positive influence is when variable X is high while variable Y is also high; negative influence is when variable X is high while variable Y is low. If the variances for each component are equal the equation is

$$Var(X + Y) = Var(X) + Var(Y) \pm 2\ COV(X, Y). \quad [13]$$

In contrast, for unequal variances for each component the equation is

$$Var(aX + bY) = a^2Var(X) + b^2Var(Y) \pm 2abCOV(X, Y). \quad [14]$$

For this case the additional variables a and b define the proportion (*i.e.*, weight) of the variance for each component for the overall sum of variance.

CHAPTER THREE
MATERIALS AND METHODS

This thesis applies continuous improvement techniques for the cellulosic biomass supply chain. A simulation model in the context of continuous improvement techniques was developed to identify components in the supply chain that are inducing the most variation. A handbook was developed for practitioners as a template for continuous improvement as part of the thesis.

Simulation Model

The success of the cellulosic biofuel production depends on the efficiency of the preprocessing technologies, conversion technology, and biomass supply. A large problem for the competitiveness of the biofuel production is the high variation associated with the quality of the supplied cellulosic feedstock. Therefore, a comprehensive simulation tool was developed to quantify the financial loss due to variation (*i.e.*, variance) in key quality characteristics of biomass feedstocks for an improved biomass supply chain, *i.e., the components of the supply chain are represented as a series system.* This technique for practitioners of the bio-based and forest products industries is also applicable outside these industries.

This research emphasizes the impact of feedstock variation in manufacturing and its influence on financial loss. This simulation tool also helps practitioners to visualize variation and identify the component inducing the most financial loss. This tool will hopefully lead to reduced variation in key quality characteristics of biomass and to a more robust product, *i.e., competitive commodity feedstock with low varying quality characteristics for increased conversion yield at biorefineries.* Figure 18 illustrates the theory of robust product design in the context of Taguchi's '*signal-to*-noise' ratio. An increase of the product quality (x-axis) leads to smaller variation of the key quality characteristics of the product (y-axis), which ultimately leads to less financial loss (Taguchi, 1993).

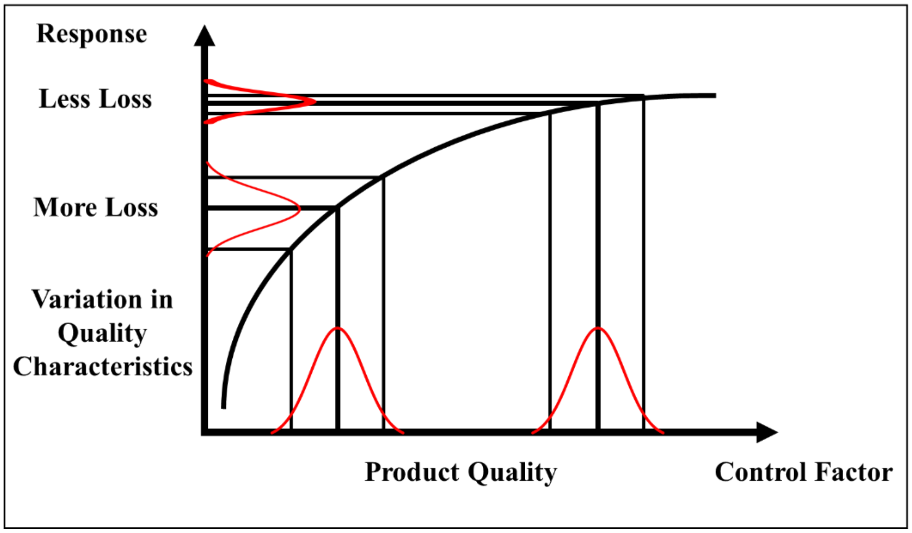


Figure 18. Taguchi Robust Product Design – increased product quality lead to reduced variation and less loss (Taguchi, 1993).

Microsoft Corporation's Excel 2016 with its integrated programming language Visual Basic of Applications (VBA) was the platform used for the simulation model. An introduction with instructions on using the spreadsheets, which include data inputs and outputs, are included in the simulation tool (Attachment File 1).

Supply Chain Design

The advanced uniform format feedstock supply chain system (Hess et al., 2009) was selected as an representative biomass supply chain, *i.e., representative series system for the simulation*. This supply chain system allows the production of standardized cellulosic feedstock products, *e.g., pellets.* The series is simplified as follows (Figure 19): 1) harvest and collection, 2) preprocessing, 3) storage, 4) transportation and handling, and 5) receiving. Preprocessing operations take place in a 'bio-depot' and consists of chipping (knife-ring flaker), drying (rotary drum dryer), blending (hammer mill), and densifying (pellet mill) the harvested biomass, also see the simulation model by Platzer (2016). The targets and specification limits were obtained from the literature (Jacobson et al., 2014, Tumuluru et al., 2014).

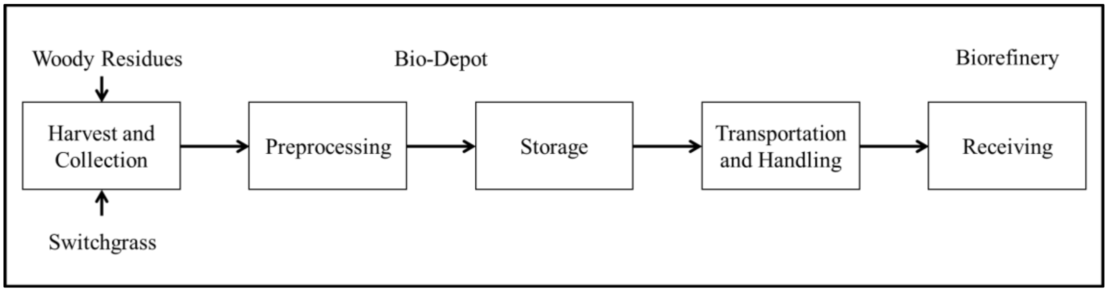


Figure 19. Simplified advanced uniform format feedstock supply chain system.

Key Feedstock Quality Characteristics

Ash content, moisture content, and particle size were selected as the key quality characteristics for cellulosic biomass conversion to biofuels based on the results of previous research (Kenney et al., 2013, Li et al., 2016, Platzer, 2016, Williams et al., 2015). Each quality characteristic impacts the performance of the supply chain and its components. Thus, visualizing the financial impact of present variation in these quality characteristics should be highly prioritized. For example, high ash content reduces biofuel yield at the conversion process (US Department of Energy, 2014), high moisture content aggravates biomass handling and transport (Eggink et al., 2018), and particle size impacts also the conversion process (Kenney et al., 2013).

Statistical Methodology

Genichi Taguchi's quality loss functions (Taguchi et al., 2004) are applied to quantify the financial loss based on variation (*i.e.*, variance is defined as σ^2) in the key quality characteristics of cellulosic feedstock. Recall from Chapter Two, Taguchi's philosophy, *i.e., monetary loss is experienced at the very moment a characteristic of interest y of a product deviates from the target m.* The loss is determined by this deviation and the proportionality constant k; k is the ratio of the maximal acceptable monetary loss (A_0) at the specification limits and the customer tolerance (Δ_0) (*i.e.*, specifications limits).

The two-sided quality loss function *nominal-the-best* is applied for the quality characteristics moisture content and particle size in equations [3] and [4] for symmetric specifications (Figure 20), and is as follows:

$$L = k \times [\sigma^2 + (\bar{y} - m)^2], \quad [3]$$

$$k = \frac{A_0}{\Delta_0^2} \quad [4]$$

Feedstock moisture is controllable by drying and it is crucial to find the optimal balance between reduction in moisture content and economic viability in drying cost. Biomass bulk density determines the efficiency for handling, transporting, and densifying processes and particle size reduction is important. However, fine particles negatively impact equipment and performance of most conversion technologies (Tumuluru et al., 2016). The total loss experienced at one component is calculated as the product of the average loss per unit times the sample size (*i.e.,* equation [3]).

The simulation tool recognizes asymmetric cases and allows for the quantification of the average loss per unit of each side of the target. An asymmetric case exists when either specification limits or the customer losses at the limits are different. Thus, for asymmetric cases, the simulation model quantifies the variation for both sides of the target individually, *i.e., treating the original dataset as two independent distributions* (Figure 21). The total losses based on this approach were compared with the total losses using equations [5] and [6] (Liao, 2010), which calculate the loss for each individual value and later summed up. The goal of this comparison is to check if the total loss based on the average losses per unit for asymmetric *nominal-the-best* cases provides a good estimate of the more precise total loss using equations [5] and [6]. The introduced approach would provide information on the average loss per unit induced by variation in quality characteristics.

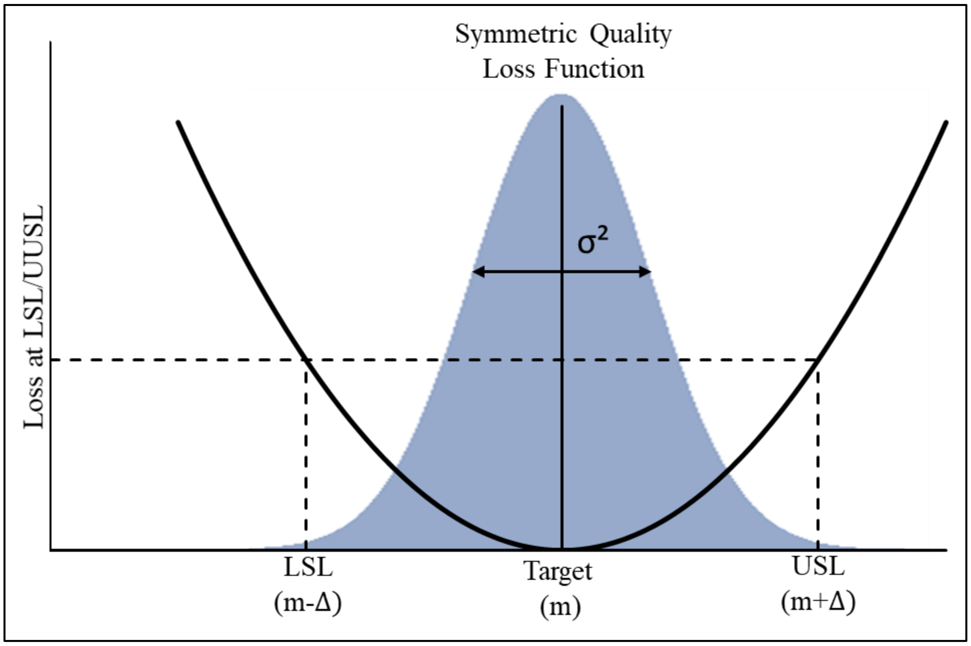


Figure 20. Schematic illustration of the symmetric two-sided quality loss function.

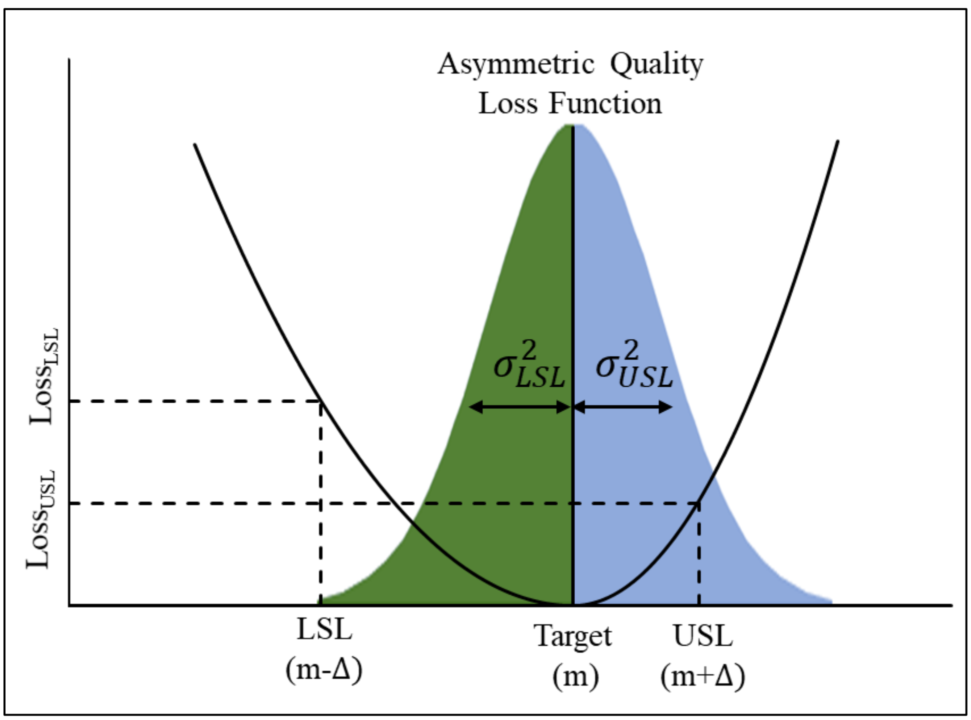


Figure 21. Schematic illustration of the asymmetric two-sided quality loss function.

The average loss per unit L_{LSL} for values below the target can be calculated as follows:

$$L_{LSL} = k_{LSL} \times [\sigma_{LSL}^2 + (\bar{y}_{LSL} - m)^2], \qquad [15]$$

with the cost constant k_{LSL} for the lower side of the two-sided loss function,

$$k_{LSL} = \frac{A_{0,LSL}}{\Delta_{0,LSL}^2} = \frac{A_{0,LSL}}{(m-LSL)^2} \qquad [16]$$

Where:

σ_{LSL}^2 = variance of all values below the target in the data set,

$\bar{y}_{LSL}$ = mean of all values below the target in the data set,

m = target,

LSL = lower specification limit,

$A_{0,LSL}$ = consumer loss at LSL,

$\Delta_{0,LSL}$ = consumer tolerance.

The average loss per unit L_{USL} for values above the target can be calculated as follows:

$$L_{USL} = k_{USL} \times [\sigma_{USL}^2 + (\bar{y}_{USL} - m)^2], \qquad [17]$$

with the cost constant k_{USL} for the upper side of the two-sided loss function.

$$k_{USL} = \frac{A_{0,USL}}{\Delta_{0,USL}^2} = \frac{A_{0,USL}}{(m-USL)^2}. \qquad [18]$$

Where:

σ_{USL}^2 = variance of all values above the target in the data set,

$\bar{y}_{USL}$ = mean of all values above the target in the data set,

m = target,

USL = upper specification limit,

$A_{0,USL}$ = consumer loss at USL,

$\Delta_{0,USL}$ = consumer tolerance.

The one-sided quality loss function *smaller-the-better* is used for computing the loss for variation in ash content; recall equations [7] and [8], Chapter Two. Ash content in biofuel feedstock decreases the biofuel production yield and therefore ash content should be small as possible, optimally, but unrealistically zero percent. The simulation tool allows the user to select the most suitable quality loss function for each component of the supply chain individually. The cost constant *k* will be modified for each individual equation based on specification limits from the literature. However, the maximum acceptable customer loss is different for each individual biomass supply chain and is not published in the literature. Values were assumed in the simulation.

By applying Taguchi's quality loss function, the simulation tool computes the financial loss based on variation (variance) for each component individually. Nevertheless, the variation of one component could have either a negative or positive effect on the actual variation of the following component and thus, change the financial loss.

Therefore, to emphasize this phenomenon the simulation model applies Galton's theory of cumulative variances for a series system. For example, assume a series system with four components. The variance, based on Galton, for the last component would be the sum of all variances and either positively or negatively impacted by twice the sum of the covariances between components. Due to lack of data in the published literature, it was not possible to use the weighted equation. Thus, the following general equation is used to calculate the variance for each step (Figure 22).

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) \pm 2 \times \sum_{1 \leq i < j \leq n} Cov(X_i, X_j) \qquad [19]$$

where:

$Var(\sum_{i=1}^{n} X_i)$ = Computed variance for n supply chain steps;

$\sum_{i=1}^{n} Var(X_i)$ = Sum variances for n supply chain steps;

$\sum_{1 \leq i < j \leq n} Cov(X_i, X_j)$ = Covariance between supply chain step i and j.

These variances are eventually used within the quality loss functions to compute the monetary loss for each component of the series. Each individual loss is added together to generate the total loss for one specific quality characteristic.

$$L_{Total,quality\ characteristic} = L_\alpha + L_\beta + L_\gamma + L_n + \cdots \qquad [20]$$

where:

$L_{Total,quality\ characteristic}$ = Total monetary loss for a certain quality characteristic (*e.g.*, ash content, moisture content, or particle size),

L_α = Monetary loss for first component in the series,

L_β = Monetary loss for second component in the series,

L_γ = Monetary loss for third component in the series,

L_n = Monetary loss for a certain quality characteristic (*e.g.*, ash content, moisture content, or particle size) at supply chain step *n*.

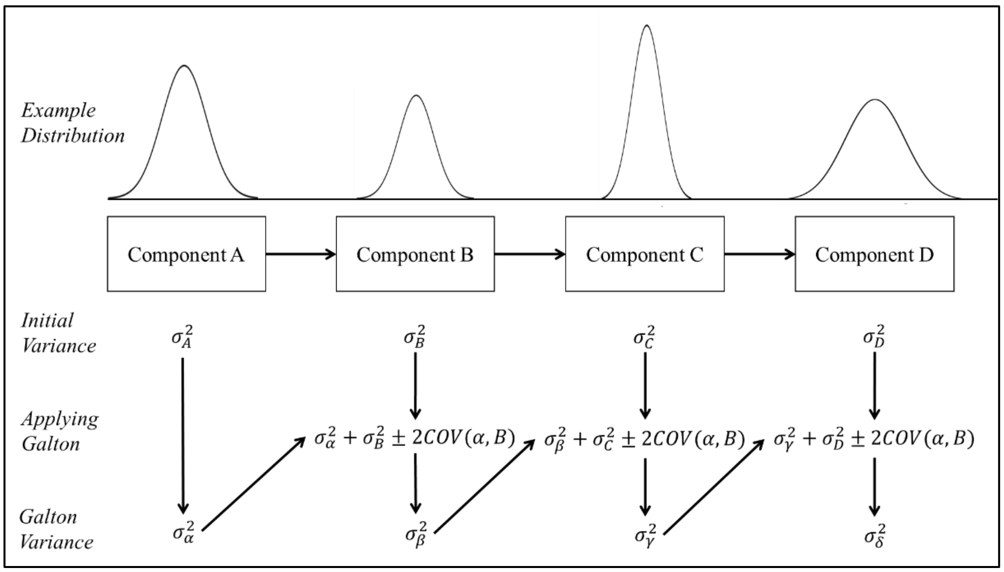


Figure 22. Example scheme for the application of Galton's theory for a series system

Materials

Data from previous research[2] for Switchgrass were analyzed with the simulation tool. However, the collected data does not provide enough information regarding ash

---

[2] The work was completed under the DOE-funded "Logistics for Enhanced-Attribute Feedstocks" (LEAF) Project, and this material is based upon work supported by the Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), under Award Number DE-EE0006639

content, moisture content, and particle size across all supply chain steps. The data introduced only provides Switchgrass samples (n = 137) for ash content at the harvesting / collection process. The Switchgrass samples were collected from several harvest sites from East Tennessee near Vonore and blended to one batch and afterwards drawn from one batch (Figure 23). In addition, simulated data were used to demonstrate the simulation tool for the series system (Attachment File 2).

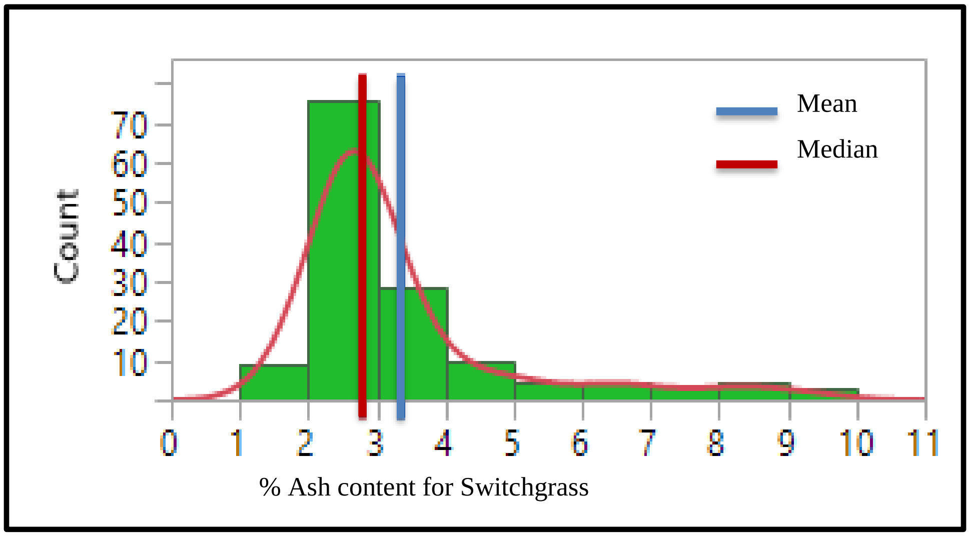


Figure 23. Histogram for ash content (%) of switchgrass (n=137) at the harvesting / collection process.

Bootstrapping was applied to calculate the necessary statistics, *i.e.*, *mean*, *variance*, *and covariance.* Statistical bootstrap is a resampling technique and uses observed data (*i.e.*, original sample) to estimate the sampling distribution (Hesterberg, 1998). For this, the observed data must be assumed to be representative of the population where it is drawn from. Starting the procedure with drawing single values from the original sample, storing them into the bootstrap sample, and eventually *put the value back* in the original sample. Values for the bootstrap sample are drawn until the size of the original sample is reached. Now, the statistics of interest (*e.g.*, mean or variance) are computed for the bootstrap sample. This procedure is done hundreds or thousands of times (Pottel, 2015), eventually creating a normal distribution for the statistics which allows to compute a grand value for each statistic. The code for the simulation was inspired by several references (Alexander and Kusleika, 2016, Verschurren, 2014).

Sensitivity analyses using empirical examples (Table 5) with real or simulated data (*e.g.*, Switchgrass, ash content) were conducted to see how variation (variance) for a given mean, target, and specification limits impacted the loss from the Taguchi Loss Function.

For a better illustration of Taguchi's quality loss functions, the computed losses (*i.e.,* average loss per unit) presented in the Results and Discussion chapter are representative for one batch of cellulosic biomass. One batch (*i.e.,* one unit) represents one dry ton of cellulosic biomass. Thus, the average loss per unit can be understood as the average loss per dry ton. Furthermore, assume that the cellulosic ethanol biorefinery with a capacity of 20 million gallons per year is able to produce on average 80 gallons of cellulosic ethanol per dry ton of cellulosic biomass. Given those assumption 250,000 dry tons of cellulosic biomass are necessary to meet the production capacity.

Continuous Improvement Handbook

As previously indicated, a handbook was developed to introduce core techniques of statistical process control and lean management to practitioners in the sustainable biomaterials industries. This handbook is a suggested template for applying these techniques. A short introduction in descriptive statistics, SPC – control charts, Taguchi's quality loss functions, and lean manufacturing procedure are included in the handbook.

Table 5. Empirical examples used for sensitivity analysis.

Example	Purpose	Components	Quality Characteristic
1	Real Example + Sensitivity Analysis	**Harvest / Collection** USL[1] = 4% Loss at USL[2] = $20	• Ash content of switchgrass • Smaller-the-better
2	Illustration of a series system for a bio-depot + Galton's theory	**Harvest / Collection** Target[1] = 40%; LSL/USL[2] = 37 / 43%; Loss at Limits[2] = $5 **Transport** Target[1] 30%; LSL/USL[2] = 27/33% Loss at Limits[2] = $15 **Drying** Target[1] 30%; LSL/USL[2] = 27/33% Loss at Limits[2] = $10 **Densification** Target[1] 19%; LSL/USL[2] = 17/21% Loss at Limits[2] = $20	• Moisture content of woody residues from harvest until end of preprocessing • Nominal-the-best • Fictitious data with n = 100 for each component
3	Asymmetry – quadratic loss function	**Densification – Cuber** Target[2] = 13.5 mm; LSL/USL[3] = 12-16 mm Loss at Limits[2] = $20	• Particle size at the densification process • Nominal-the-best • Fictitious data with n = 100

[1] Taken from Jacobson et al. (2014)
[2] Assumptions made for the simulation
[3] Taken from Tumuluru et al. (2011)

CHAPTER FOUR
RESULTS AND DISCUSSION

Variation in product quality characteristics is a key factor in limiting the technological and economic performance of biomass supply chain operations and biomass-to-biofuel conversion technologies. Thus, controlling and reducing the underlying variation of core quality characteristics was studied in this research, *e.g., ash content, moisture content, and particle size variation reduction has promising potential to increase the viability of sustainable bio-based productions.*

For example, increasing attention towards improving preprocessing technologies and supply chain design concepts (Hess et al., 2009, Platzer, 2016) allow more efficient supply of standardized feedstocks while simultaneously meeting the established technological requirements of the biorefinery. Therefore, visualizing and quantifying variation across supply chain operation units or production process units offer great incentives to act and provide a solid foundation for managers to optimize their productions.

As part of this thesis an Excel simulation tool (available at www.spc4lean.com) was developed to quantify the financial loss through feedstock variation for a simplified series system, *e.g., cellulosic biomass supply chain*. The main goal of this simulation tool is to expose practitioners to the effects of variation on the financial loss as exemplified by Taguchi's quality philosophy (Taguchi et al., 2004). Empirical examples are given for estimating loss. An instructional handbook outlining continuous improvement techniques, SPC, and Taguchi's philosophy was developed to provide a template for practitioners of the sustainable bio-based industries for the improvement of production systems.

A Guide for Using the Simulation Tool

Spreadsheet 1 – Content

The Excel workbook starts with the spreadsheet labeled *'Content'* (Figure 30), which provides an overview of all included spreadsheets. Each spreadsheet can be accessed through a bold and underlined hyperlink. Cell A1 of every spreadsheet contains a hyperlink called “Content”, which leads back to the content page. The workbook consists of an introduction to the topic and a help guide for the simulation. Furthermore, a bootstrap simulation for non-parametric data to compute the financial loss based on variation of quality characteristics forms the main part of the simulation. Further data analysis can be done on the spreadsheets *‘Sensitivity Analysis’* and *‘Galton Theory’*. The spreadsheet *‘Computations’* provides the results of auxiliary computations of the mean, variance, or covariance of each component of the series. The workbook is concluded with a *‘Summary’* of the simulation output.

Spreadsheet 2 – Introduction and Help Guide

The second sheet labeled *‘Introduction and Help Guide’* (Figure 31) introduces the user to the advanced uniform feedstock supply system, traditional quality control, Taguchi’s quality loss function philosophy, Galton’s theory, bootstrapping, and a help guide for the simulation. The main simulation consists out of two parts which are the following: User input and simulation output.

Spreadsheet 3 – User Input

Spreadsheet three labeled *‘User Input’* (Figure 32) provides the environment for the user to set the parameters for the bootstrap simulation and to enter necessary variables for each component of the series system. The sheet is structured in two parts. A help guide, the first part, placed on the left on the sheet, consisting out of six steps and will help the user navigate through the preparation process of the simulation. The empty space on the right is reserved for an input-table created at step three of the help guide, which builds the second part of the sheet.

The help guide starts with the introduction of key quality characteristics for the cellulosic biomass supply chain and their respective quality loss functions, as well as required input about general information of the analyzed quality characteristic, *i.e., name (ash content), unit of measurement (%), and currency ($).* The next step determines the number of components for the series system; up to 12 components can be analyzed at a time. The input-table is created by clicking on the button *'Make Table Design'*. The first column of the table indicates a set of key variables, which must be provided by the user for each component of the series. For greater individualization and a better reflection of reality the user can independently select the quality loss function type (*e.g., nominal-the-best, smaller-the-better*, and *larger-the-better)* for each component. Depending on the selection several system related variables must be entered to run the simulation.

For *nominal-the-best* cases values for the target, upper and lower specification limit, loss at upper and lower specification limit must be provided by the user. The simulation allows the entry of *symmetric* and *asymmetric* specification limits. Taguchi provides the equation [3] for calculating the loss based on the variance for symmetric cases. For asymmetric cases the simulation computes the approximated loss for the data of either side of the target (recall Figure 21) based on the variance and mean; a closer discussion can be found in a later section of this chapter. For *smaller-the-better* and *larger-the-better* cases the user only needs to input values for the target and the loss at target. Afterwards, starting with cell J13 in the spreadsheet the user should enter the measured values of their quality characteristic for each component. The hard-coded maximum size of a data set is 5,000 values.

After setting up the input table and entering all necessary parameters and values, the number of iterations for the bootstrap simulation needs to be entered. Usually 5,000 to 10,000 or even more iterations are done to generate statistically acceptable results.

The goal of this bootstrap simulation is to simulate different "collected/measured" sets of data to generate a range of values of a certain statistic, *e.g., mean, variance, and covariances of all components,* to find the grand values of these statistics. Additionally, the number of bins for the histograms of the statistics on sheet four *Simulation Output* can be set by the user. The final step of the data input phase is to click on the button *'Execute Simulation'* to run the simulation. Primarily, the financial losses are computed as a total and as the average per unit based on Taguchi's quality loss functions. However, the simulation will provide these losses computed with two different variances, *i.e., independent components in a series system and dependent components based on Galton's* *theory of components of variances*. The code for the bootstrap simulation can be found in appendix A.

These resulting losses based on Galton show how the variation of components in the series system impact each other. Only simulation results for *nominal-the-best* and *smaller-the-better* cases provide information about the impact of component dependency. A summary of simulation is given with the creation of several charts and histograms.

Due to limited information from the literature, the main simulation from sheet three uses the unweighted equation and treats each components variance as equal, *i.e., each component has the same impact on the final loss of the system.*

Spreadsheet 4 – Simulation Output

Spreadsheet four labeled *'Simulation Output'* (Figures 33 and 34) presents the computational and graphical output of the data analysis based on the user input from sheet three *'User Input'* for each component of the series system. The orange colored area is divided into two sections. The first section provides input values such as target, USL, LSL, loss at USL or LSL, position and name of the component, as well as which quality loss function type was used. The second part refers to the simulation output the computational results such as constant *k*, average loss per unit and total losses based on the variance for independent components and dependent components (Galton's theory) in a series system. The total loss per component is computed as the product of average loss per unit times the number of the initial values of the original sample given on sheet three. The total loss of the series system is displayed on the left and is the sum of the total losses of all components. Below the computational output graphical displays of the quality loss function for each component are shown. The first and second chart differs solely based on the data distribution curve. The first chart shows the data distribution for the original sample, while the second chart shows all values drawn within the bootstrap procedure. This allows an interesting comparison between the real initial data distribution and the simulated data distribution. The red graph emphasizes the quality loss function (first y-axis on the left) and the lavender blue graph represents either data distribution. The grey bar stands for the target, the green bar highlights the lower specification limit, and the purple bar emphasizes the upper specification limit. The second y-axis on the right stands for the number of quality characteristic values (x-axis) of the original sample or drawn by the bootstrap procedure.

The simulation also produces several histograms. For *nominal-the-best* and *smaller-the-better* cases the bootstrap distributions for the mean and the variances are shown, as well as the grand-values of these statistics. Due to increased comparability the bootstrap distributions of either statistic for both sides of an *asymmetric nominal-the-best* loss function are each shown in one histogram.

Spreadsheet 5 – Sensitivity Analysis

Spreadsheet five labeled *'Sensitivity Analysis'* (Figure 35) provides a sensitivity analysis tool to estimate the average loss per unit of any component of the series relying on the bootstrap simulation for estimates of the parameters. The sheet is structured in two parts; a help guide on the right and an overview of the in- and output values on the left. The first step is to run the bootstrap simulation from spreadsheet three to generate the required input data for the sensitivity analysis tool. Step two is the selection of the component of the series the user desires to analyze. By clicking on the button *'Load Data'* the initial input values and the results from the simulation are shown on the left. The embedded VBA code recognizes the chosen quality loss function type for the selected component; the sheet automatically adjusts the output based on the type. For example, user input such as target or specification limits and computed results like the mean, variance, and losses are displayed to provide an overview of the respective component. The actual sensitivity analysis takes place through step four and five, which can be executed as many times as wanted. In addition to the loaded values on the left, the spreadsheet shows a variety of changeable variables on the right under step four. This feature represents the actual sensitivity analysis, *i.e., the user can enter a value for any variable and compute the loss.*

For example, in *nominal-the-best* cases, variables such as target, upper and lower specification, loss at these specification limits, mean, and variance. The embedded VBA code is sensitive to the given input, *i.e., the code checks whether cell is filled with a value.*

By default, the code uses the initial values from the bootstrap simulation. However, if the user entered a new and different value for any variable in step four, the code computes the loss with this new value. This feature is enabled for all different quality loss function types. Nevertheless, for asymmetric cases the user must specifically decide which side of the quality loss function he wants to investigate by checking the dropdown list. The newly computed average loss per unit is shown in an orange box. Above the orange box the respective constant *k* is shown as well. This feature allows the user to analyze and see the effects of variation on the financial loss. As an optional feature, the user can click on the button *'Save the Data'* to save the computed losses in a table. This allows the user to create a table for sensitivity analysis. For each click on the button an internal count is incremented by one to move to the next row to avoid overwriting of values.

The simulation can be restarted for a new set of data for a component of the series system by clicking on the button *'Reset Saved Data'* the whole table is cleared and the count is reduced to one again. Depending on the quality loss function the table includes the computed loss from the orange box, constant *k*, variance, and mean. However, since the *larger-the-better* loss functions does not use either the mean or the variance to compute the average loss the changeable variables are only target and loss at target.

Spreadsheet 6 – Galton Theory

Spreadsheet six titled *'Galton's Theory'* (Figure 36) analyzes Sir Francis Galton's theory of cumulative variances for a series system; recall equations [13] and [14], see (Stigler, 2010). The total variance for a series system is the sum of all variances and twice the sum of all covariances between all components of the system. All variances of the system are assumed equal in equation [13]. In reality, variances are likely to be unequal and specific weights for each component will be included in the calculation. These weights show the *true* impact of a components variance on the total system's variance. Estimated model coefficients based on a multiple linear regression (MLR) analysis are used as weights for the loss computation. Since MLR only provides coefficients for the explanatory variables, the weight of one is used to explain the impact of the variation at the final stage of the system; the response variable. Coefficients of the MLR independent variables are the weights, *i.e., the total variation as a percent is 100%.* The goal of this sheet is to compare the approximated average loss per unit computed with independent variance (Galton's unweighted equation) with Galton's weighted equation for a series system.

The spreadsheet is structured as the following: On the left side the user can enter data for up to 12 components of the series. By selecting either *yes* or *no* for each component the embedded VBA code recognizes the component selected for the MLR. Importantly, the last component filled with data functions as the response variable, *e.g., five components are selected, the fifth component or column represents the data for the response variable.* All other components function as explanatory variables. As mentioned model coefficients are used to determine each component variations' impact on the total system variance. In addition, values for the variables constant *k* and target must be entered for each component. After entering all values and finished selection, press button *'Compute Loss for a Series'* to execute the MLR procedure. The embedded VBA code uses the MLR procedure from Microsoft Excel; the add-in *'Analysis Toolpak VBA'* must be enabled. The created output includes the MLR output, the covariance matrix, and several computed statistics. At the top of the sheet the output for the computed losses for different variances (comparing Galton's equal variances with Galton's unequal variances) are presented for both the *nominal-the-best* or the *smaller-the-better* loss functions.

Spreadsheet 7 – Summary

The final spreadsheet titled *'Summary'* (Figure 37) shows the main results from sheet four. The loss for each component and the total loss for the whole series system are displayed. This summary is in keeping with the theme of the thesis, *i.e., to emphasize to the practitioner the effect of variation and cost due to variation and components of variation in the process.*

Numerical Examples and Sensitivity Analysis

Empirical Examples

The following empirical examples were developed to highlight the capability of the simulation tool. Losses in the empirical study are for the biomass supply chain and its components for *nominal-the-best* (*e.g.,* moisture content and particle size) or *smaller-the-better* (*e.g.,* ash content) quality characteristics. The assumption in the empirical study was that the data sets for each component follow a non-parametric distribution using the bootstrap to simulate the distribution of key statistics (N = 10,000 iterations). Effects of changes in variance and/or shifts in mean are presented in the sensitivity analyses.

Example 1 illustrates Taguchi's *smaller-the-better* loss function for varying ash content of Switchgrass at the harvesting and collection operation for the biomass supply chain.

Example 2 analyzes the effects of variation in moisture content of woody residues on the loss for Taguchi's *nominal-the-best* loss function for a simplified biomass supply chain and its components. Woody residues are collected and transported to the *'bio-depot'* for further preprocessing to generate in quality characteristics standardized products. At the bio-depot the woody residues are dried and densified to pellets.

Example 3 demonstrates the use of Taguchi's quality loss function to compute the total loss based on the average loss per unit (*i.e.,* equation [3]) for a component for *nominal-the-best* quality characteristics in *symmetric* and *asymmetric* specification settings and is compared to the total loss computed with the sum of the individual losses. Data for particle size of cellulosic biomass at the densification process using a *cuber* is used to illustrate this example.

Example 1 – Loss Caused by Variation in Ash Content of Harvested Switchgrass

The developed Excel spreadsheet was used to generate the bootstrap simulation and generated losses, bootstrap statistics, and graphical outputs for the ash content of harvested Switchgrass. The mean value from the bootstrap simulation is $\bar{x}$ = 3.35% and the variance is σ^2 = 2.68%². The parameters vary little compared with the mean $\bar{x}$ = 3.35% and variance

σ^2 = 2.70%² for the original sample. The *smaller-the-better* loss function is illustrated in Figure 24 (*i.e.,* Equation [7]) for ash content in red exponentially rising for greater ash content values. The histogram in green represents the distribution of the original sample data. Most of the data are below the USL shown as a black line of four percent and are close to the mean depicted as a blue line. However, many values are outside of the specification limit, *i.e., above the USL.* The financial loss for *'out-of-spec'* data is significantly higher than for data within specification limits. In comparison, the smoothed data distribution in green for the drawn samples within the bootstrap simulation is shown in Figure 25 and shows a peak around three percent ash content and is very similar to the original sample distribution. Figures 24 and 25 illustrate how Taguchi's *smaller-the-better* quality loss function works and how extreme variation in quality characteristics impact the final loss. For example, the loss for supplied Switchgrass with an average ash content of three percent per batch (*i.e.,* dry ton) is $11.25 with a cost constant *k of* $ 1.25 $(\%^2)^{-1}$. Accumulating in an annual loss for the assumed biorefinery with a demand of 250,000 dry tons of cellulosic biomass of $2.8 million dollars. Now for a doubled ash content value (6%) the average loss per dry ton would be $45. Increasing the annual loss, induced through feedstock variation, for the same biorefinery by $8.5 million to $11.3 million dollars. Both examples illustrate how dramatic the loss for *smaller-the-better* quality characteristics increases for higher variations. Engineers and manager should see the loss as an indicator for the component in the production system which has the highest impact on the economic performance.

The average loss per unit (*i.e.,* average loss per dry ton) based on Taguchi's *smaller-the-better* loss function (*i.e.,* equation [8]) for the ash content of harvested Switchgrass computed with the bootstrap statistics and a cost constant *k* of $ 1.25 $(\%^2)^{-1}$ is $17.37 per dry ton. The average loss per unit is high due to the skewness of the ash content distribution, *i.e., several Switchgrass samples had a higher ash content than the selected upper specification limit (USL) of 4%.* The annual total loss for the assumed biorefinery would then be $4.3 million dollars, *i.e., demand of 250,000 dry tons cellulosic biomass times $17.37 per dry ton*. The total loss assumes that the ash content of Switchgrass was not reduced by preprocessing to increase the conversion yield.

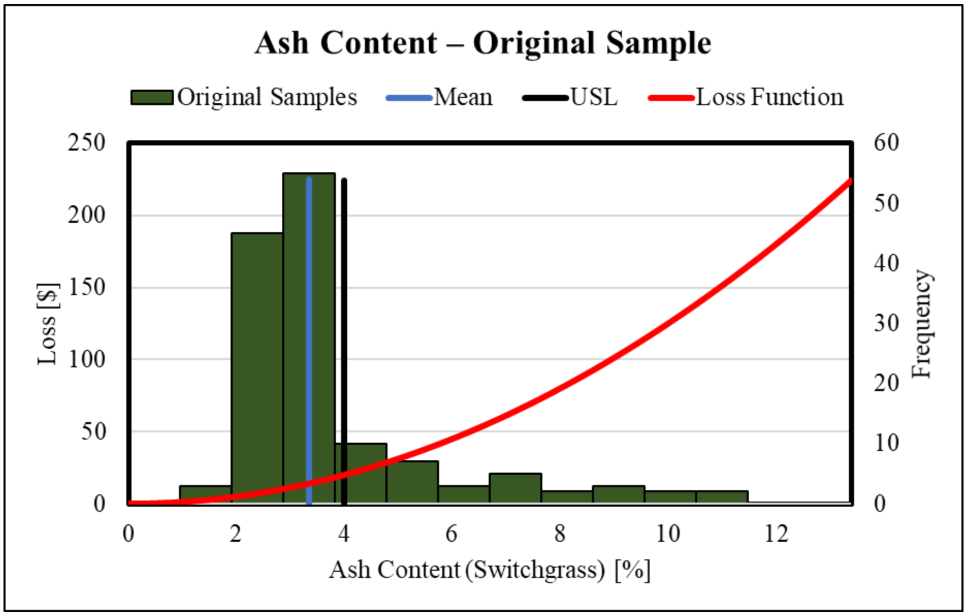


Figure 24. *Smaller-the-better* loss function for the ash content of harvested Switchgrass for the original sample. Equation [7] is used to generate the loss function, *e.g.*, $L(4\%) = \$\ 1.25\ (\%^2)^{-1} * (4\%)^2 = \$20$.

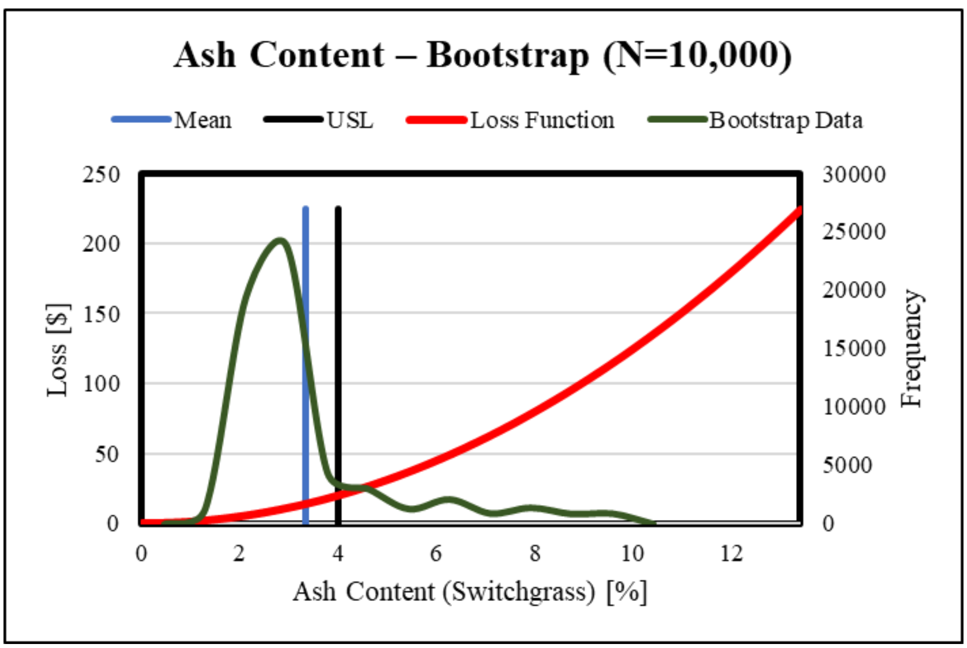


Figure 25. *Smaller-the-better* loss function for the ash content of harvested Switchgrass for the bootstrap data. Equation [7] is used to generate the loss function, *e.g.*, $L(4\%) = \$\ 1.25\ (\%^2)^{-1} * (4\%)^2 = \$20$.

A sensitivity analysis was conducted to illustrate the patterns of the average loss per unit for the *smaller-the-better* quality loss function based on changes in the mean, variance, USL, and customer loss at USL (Table 6). Each parameter was either increased or decreased by 0.1 from their respective original value used to calculate the average loss per unit of $17.37, while the other parameters were kept constant. The sensitivity analysis indicates that the average loss per unit for ash content is more sensitive to small changes in the mean and in the USL than to changes in the variance and customer loss at USL. Due to the design of Taguchi's *smaller-the-better* loss function the average loss per unit increases for reduced USL, *i.e., the cost constant k increases due to the USL value being the squared denominator for unchanged customer loss at USL, recall equation [4]*. Furthermore, the cost constant *k* shows greater increase and decrease in changes in the USL than in the loss at USL.

Additionally, the average loss per unit increases the closer the mean gets to the USL, *i.e., the average value of all ash content samples of the batch deviates further from, what Taguchi's smaller-the-better loss function defines as the optimal target, zero.* However, running zero is theoretical as an operational target not achievable. For example, it is very difficult to achieve cellulosic feedstocks with zero percent ash content to increase the conversion yield at the biorefinery, *i.e., cellulosic biomass possesses structural ash content within their cells* (Lacey et al., 2016). Cost-intensive pretreatment and optimized harvesting schemes would allow for a reduction of the ash content, but these efforts to decrease variation and to move the mean closer to the optimal target, to increase the conversion yield of ethanol, may not be economically justifiable. However, prescreening of feedstock vendors may be helpful in identifying those vendors that have the largest ash content means and variance. The cost constant *k*, computed as the quotient of customer loss at USL and USL squared, would be smaller because the deviation of the USL from zero is greater than the deviation of the USL from any value greater than zero. Nevertheless, it is more realistic given that zero ash content is not attainable.

Figure 26 illustrates the average loss per unit (y-axis) for each changed parameter (x-axis) as a line. The average loss per unit follows a quadratic pattern for shifts in the mean and USL, depicted by green and blue vertical lines, respectively. Changes to variance

Table 6. Sensitivity analysis of ash content for harvested Switchgrass *smaller-the-better* loss function.

k	USL in [%]	Loss at USL in [$]	Mean in [%]	Variance in [%²]	Loss per Unit in [$]
1.39	3.80				19.24
1.31	3.90				18.27
1.25	4.00	20	3.349	2.680	17.37
1.19	4.10				16.53
1.13	4.20				15.75
1.24		19.80			17.19
1.24		19.90			17.28
1.25	4.0	20,00	3.349	2.680	17.37
1.26		20.10			17.45
1.26		20.20			17.54
			3.149		15.74
			3.249		16.54
1.25	4.0	20	3.349	2.680	17.37
			3.449		18.21
			3.549		19.09
				2.480	17.12
				2.580	17.24
1.25	4.0	20	3.349	2.680	17.37
				2.780	17.49
				2.880	17.62

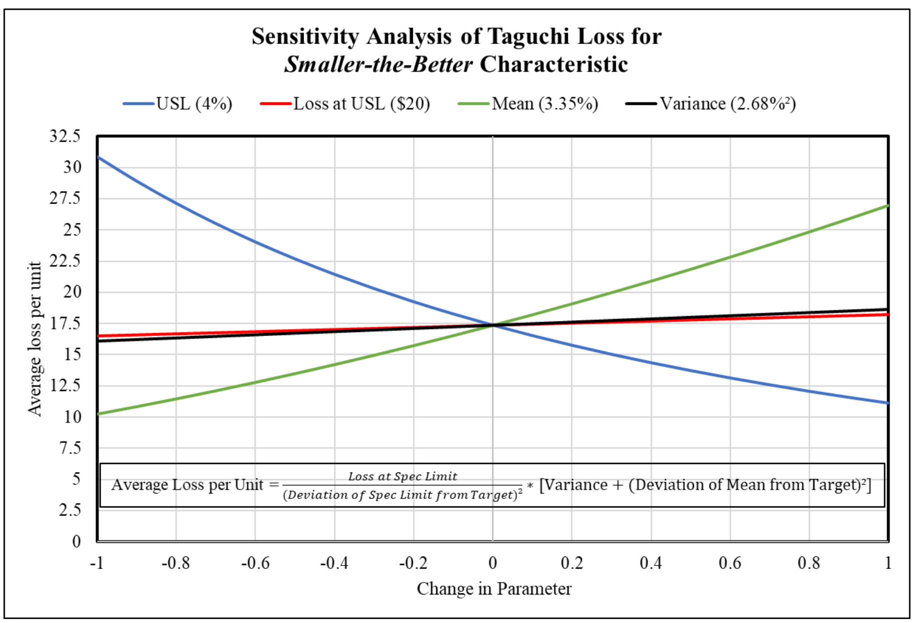


Figure 26. Sensitivity analysis of the average loss per unit of Taguchi's quality loss function *smaller-the-better* for the parameters USL, loss at USL, mean, and variance.

illustrated as a black line and the customer loss at USL as a red line are linear. The sensitivity analysis further indicated that the average loss per unit increases more drastically for reduced USL values compared with the average loss per unit for higher USL. This is depicted in Figure 26 by a steeper slope for a low USL and high average loss per unit versus a more gradual slope for a high USL low average loss per unit. A smaller operational upper specification limit or target imply reducing variation and a smaller natural tolerance, *i.e., higher natural tolerances impede target size reduction and does not reduce cost.* Figure 26 also suggests that the changes in the variance are less drastic to the average loss per unit than to shifts in the mean. This implies that shifts in the mean to a lower value would achieve large cost savings. However, reducing the mean is only technical possible for reduced variation. Therefore, to achieve large cost savings variation of the quality characteristic must be reduced first to be able to shift the mean closer to zero.

Example 2 – Biomass Supply Chain

The bootstrap simulation generated the bootstrap statistics and losses for a series system with four components using simulated moisture content data. The selected specification settings are suitable for woody residues. The generation of the bootstrap statistics took two minutes for the Excel simulation using an Intel® Core™ i5-4300 M CPU @ 2,60 GHz with 16,0 GB RAM. The series system consists out of representative biomass supply chain operations such as harvesting/collection, transport, drying, and densification.

Table 7 presents the means and variances for the original samples and the bootstrap statistics, as well as the coefficient of variation of the four components. The coefficient of variation for all components indicate a low variability around the mean values ranging from 4.79% for *harvest/collection* to 7.42% for *densification.* The mean values for the bootstrap statistics for all components of the series are not different from the respective original sample means. In contrast, the variance values based on the bootstrap simulation slightly differ from the variances of the original sample for all four components. For example, the variance of the original sample for the *harvest/collection* component is $\sigma^2 = 3.67\%^2$ while the bootstrap variance is $\sigma^2 = 3.63\%^2$. These results suggest that no bias for the mean values and a small bias for the variance values from the bootstrap simulation exist. The dataset for *harvest/collection* has the highest variance (σ^2 = 3.63%²) and the smallest mean deviation (0.02%) from the target. In contrast, the smallest variance (σ^2 = 2.12%²) and the largest mean deviation from the target (σ^2 = 0.73%) is experienced at the *densification* component (Table 7).

Table 7. Comparison of statistics for original sample and bootstrap data for *nominal-the-best* quality characteristics moisture content of woody residues for a simplified biomass supply chain.

Component	Harvest/ Collection		Transport		Drying		Densification	
Coefficient of Variation in [%]	4.79		5.72		5.67		7.42	
Bootstrap Statistic	No	Yes	No	Yes	No	Yes	No	Yes
Mean in [%]	40.015	40.015	30.123	30.123	29.817	29.815	19.726	19.726
Variance in [%²]	3.668	3.631	2.972	2.945	2.858	2.835	2.141	2.118

The average losses per unit computed with the bootstrap statistics for different cost constants *k* for all components of the series system are shown in Table 8. The highest average loss per unit \$13.23 is experienced at the *densification* component due to the high cost constant k = \$5 $(\%^2)^{-1}$. The other average losses per unit are \$4.93 (k = \$1.67 $(\%^2)^{-1}$) for *transport*, \$3.19 (k = \$1.11 $(\%^2)^{-1}$) for *drying*, and \$2.02 (k = \$0.556 $(\%^2)^{-1}$) for *harvest/collection* (Table 8). These results suggest that the cost constant *k* is a big driver in the average loss per unit. The annual total losses for the supply chain based on the supply for the assumed biorefinery would be \$5.8 million dollars. The individual annual total losses per supply chain operation are the following: *harvest/collection* with \$505,000, *transport* with \$797,500, *drying* with \$1.2 million dollars, and *densification* with \$3,3 million dollars. However, to better understand variation in form of variance or mean deviation the average losses per unit for all components was computed with equal *k* (Table 8). For equal cost constants the highest average loss per unit is experienced at the *harvest/collection* component. For example, the average loss per unit for the *harvest/collection* with k = \$2 $(\%^2)^{-1}$, mean of 40.02%, and a variance of 3.63%² is \$7.26. The lowest loss exists at *densification* with \$5.29 with a mean of 19.73% and a variance of 2.12%. The findings suggest for the given output a higher influence of the variance on the average loss per unit than the mean. Since equation [3] computes the average loss per unit by adding the variance to the squared difference of the mean from the target. Because this difference for all components is below one the squared results are even smaller, thus the variance has a greater influence on the average loss. Thus, implying that for *nominal-the-best* quality loss functions to save money reducing variation is imperative as it is thesis of this research. *Recall* for *smaller-the-better* loss function only the mean, not the difference, is considered, *i.e., the mean represents the difference from the theoretical desired target zero.*

Table 8. Average loss per unit using for different cost constants *k* bootstrap statistics for *nominal-the-best* quality characteristic moisture content of woody residues for a simplified biomass supply chain.

Component	Harvest/ Collection	Transport	Drying	Densification
k in [\$/%²]	0.556	1.667	1.112	5
Average losses per unit in [\$] for different cost constants *k*				
Original *k*	2.02	4.93	3.19	13.23
k = 2 \$/%²	7.26	5.92	5.74	5.29
k = 5 \$/%²	18.16	14.80	14.35	13.23
k = 10 \$/%²	73.19	63.27	50.81	68.76

A sensitivity analysis was conducted to illustrate the patterns of the average loss per unit for Taguchi's *nominal-the-better* loss function (equation [3]) for changes in the mean and the variance. Figures 27 and 28 depict these changes in average loss per unit (y-axis) in terms of six sigma (x-axis) as continuous graphs, *i.e., one sigma represents one standard deviation.* The standard deviations of the grand-variances were computed based on the spread from all individual bootstrap sample variances *(N* = 10,000) and for the grand-means are the respective square root of the grand-variances (Table 9).

Figure 27 illustrates the quadratic trend of the average loss per unit (y-axis) for the four components for continuous shift in the mean for *nominal-the-best* loss function. Generally, for a mean exact on target the lowest average loss per unit depend on the variance of the data. As previously indicated, the highest loss for minus six and plus six standard deviation (x-axis) is at *harvest/collection* (blue line). Furthermore, *drying* shown as a red line experience a higher loss at minus six sigma compared with *transport* depicted as a green line, *transport* has a higher average loss per unit at plus six sigma. This is of interest because both lines have a target of 30% and thus, the mean and variance can directly be compared. The reason for this is simply, that the mean for *drying* is below the target and the mean for *transport* is above the target (Table 7). Thus, the average loss per unit for *transport* first gets smaller the closer the mean gets to the target. Table 16 (Appendix B) presents values for average losses per unit for an incremental change of the mean in terms of six sigma for a cost constant k = \$2 $(\%^2)^{-1}$, constant variances and targets.

Table 9. Standard deviations of the bootstrap grand-mean and grand-variance for the series components.

Component	Harvest/ Collection	Transport	Drying	Densification
Standard Deviation Mean in [%]	1.906	1.716	1.684	1.455
Standard Deviation Variance in [%²]	0.470	0.323	0.349	0.427

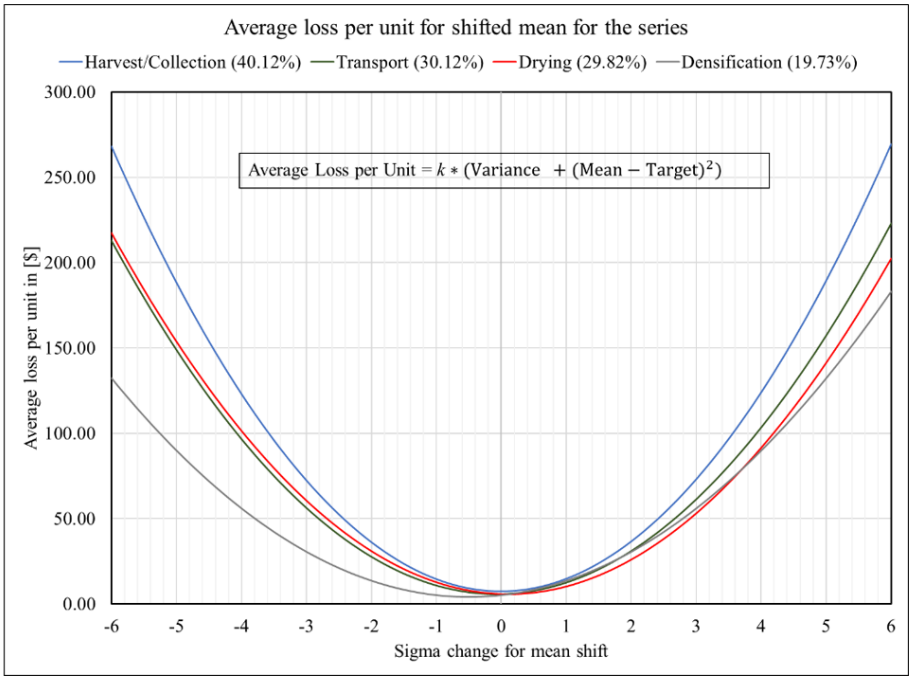


Figure 27. Quadratic pattern of the average loss per unit for *nominal-the-best* loss function for continuous shifts in mean. Losses computed with bootstrap statistics and an equal k = \$2 $(\%^2)^{-1}$.

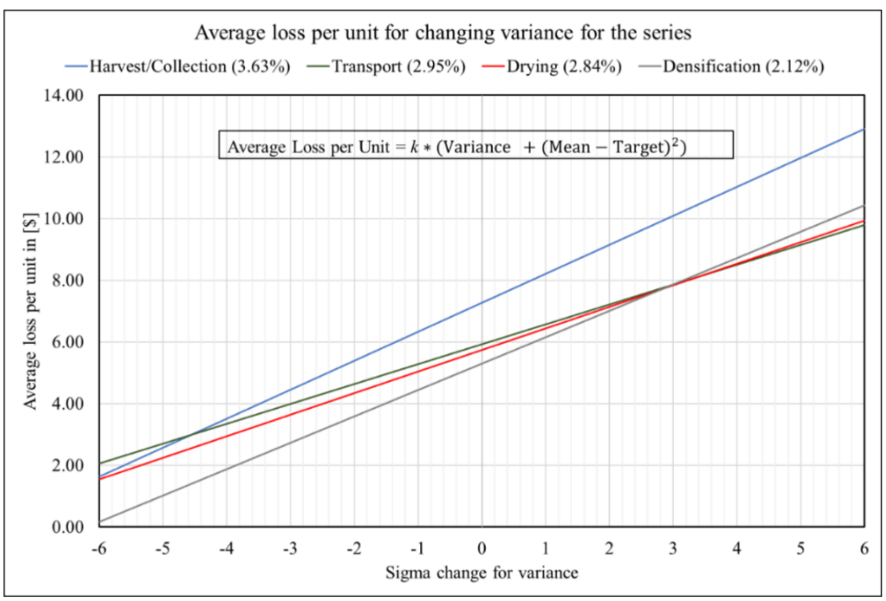


Figure 28. Linear pattern of the average loss per unit for *nominal-the-best* loss function for continuous changes in the variance. Losses computed with bootstrap statistics and an equal k = \$2 $(\%^2)^{-1}$.

Furthermore, for changes of the variance the average loss per unit based on *nominal-the-best loss* function follows a linear trend (Figure 28). As depicted as a blue line, *harvest/collection* shows the highest average loss per unit for increased variances and the steepest slope. In contrast, *densification* (grey line) has the smallest loss for decreasing variance values but has the second highest loss for variance at the positive six sigma level. The lines of *densification*, *transport* depicted as a green line, and as a red line *drying* meet slightly below the positive three sigma level. *Transport* and *drying* have a more gradual slope. The slope for the four graphs depends on the standard deviation of the variance, *i.e., the smaller the standard deviation the more gradual the slope is.* Table 17 (Appendix B) presents the average losses per unit for an incremental change of the variance in terms of six sigma for a cost constant k = \$2 $(\%^2)^{-1}$, constant variances and targets.

Application of Galton's Theory – Variance is Cumulative

The Excel spreadsheet titled *'Galton's Theory'* generated the average losses per unit to investigate the influence of the variance based on Galton's theory of cumulative variances. This influence on the average loss per unit for *nominal-the-best* loss function was computed with the statistics from the original samples (Table 7) and their respective original cost constants k (Table 8). The average loss per unit for the following cases were compared. Case 1) Series system with independent components (equation [3] without Galton), *i.e., variance is treated as non-cumulative.* Case 2) Series system with dependent components but with equal variances (equation [3] for the loss and [13] for the variance), *i.e., the variance of each component has the same impact on the total variance.* Case 3) Series system with dependent components with unequal variances (equation [3] for the loss and [14] for the variance), *i.e., the variance of each component has a different impact (weight) on the total variance.* These weights are model coefficients and stem from a multiple linear regression model applied to the data of the series conducted with Excel.

The average loss per unit for component one, for all three cases, is \$2.04 with a variance of 3.67%² (Table 10). The losses are the same for all cases because Galton's Theory is not applied for just one component of a series, *i.e., no other components impact* *the total systems' variance.* Despite having similar mean and variance values (Table 7) the average loss per unit for the components *transport* is $4.98 compared with the loss for *drying* of $3.21 is higher due to different cost constants *k* (Table 8). However, these losses change when Galton's theory is applied to compute the variances. For case two, variances are equal, the average losses increase to $9.18 for *transport* with a variance of 6.64%² and $9.11 for *drying* with a variance of 9.50%². The loss for *drying* is high since the variances and covariances of components one to three are summed up without weights (Table 11), recall equation [13]. For case three, variances are unequal, the average losses decrease to $0.21 with a variance of 0.11%² for *transport* and $0.54 with a variance of 0.289%² for *drying.* The low losses are explained by the small weights, *i.e., the variances for the components in case three are multiplied with squared model coefficients (Table 12). Densification* shows the highest average losses per unit throughout all three cases which are the following: $13.34 case one, $48.30 case two, and $14.75 case three. Due to the assumptions and design of the embedded code in the Excel spreadsheet the loss for case three is high, since a weight of one was assumed. The highest average loss per unit is experienced at *densification* for case two with $48.30. This value is based on the accumulated variance values of all four components resulting in a variance for component four of σ^2 = 9.13%². However, the losses of each component in Table 10 were computed with different cost constant *k* (Table 8). For an industrial application of Taguchi's quality loss function in combination with Galton's theory each specification limit and loss at the limit should be identified.

Tables 11 and 12 provide an overview of the composition of the calculated total variances¸ *i.e., variances used to calculate the average loss per unit under application of Galton's theory consist out of sum of variances of all components and sum of all covariances.* Table 11 shows the *sum of variance* of each component (row 2). For this value the variances from Table 5 are summed up without weights. Each components' total variance is then impacted by the sum of all covariances, *i.e., for drying covariances would exist between the first three components*. For this case the negative covariances reduce the total variances and thus reduce the average loss per unit for each component. Table 12 shows the total variance for case three. The model coefficients for *harvest/collection*

Table 10. Average loss per unit for *nominal-the-best* loss function for application of Galton's theory.

Component	Harvest/ Collection	Transport	Drying	Densification
Case	Average loss per unit in [$]			
Independent Components	2.04	4.98	3.21	13.34
Galton for equal variance	2.04	9.18	9.11	48.30
Galton for unequal variance	2.04	0.21	0.54	14.75

Table 11. Breakdown of 'Galton variance' for a series system with equal variances.

Component	Harvest/ Collection	Transport	Drying	Densification
Total Variance	3.668	5.490	8.165	9.133
Sum of Variance	3.668	6.639	9.497	11.638
Doubled sum of Covariance		-1.149	-1.332	-2.505

Table 12. Breakdown of 'Galton variance' for a series system with unequal variances.

Component	Harvest/ Collection	Transport	Drying	Densification
Total variance	3.668	0.110	0.289	2.422
Sum of variance	3.668	0.075	0.183	2.141
Doubled sum of Covariance		0.014	-0.004	-0.008
Model coefficient	-0.0764	0.159	-0.253	

(-0.0764), *transport* (0.159), and *drying* (-0.253) are very low. These low coefficients (*i.e.,* weights) reduce each components variance and covariance. This indicates that each components' variance has a small impact on the total systems variance.

Example 3 – Quantifying Loss in Terms of Variation for Asymmetric Nominal-the-Best

The Excel spreadsheet was used to generate bootstrap statistics and Taguchi losses using the same dataset of the particle size for woody residues for symmetric and asymmetric cases. The output of the symmetric and asymmetric specification settings was used to compare the total losses and average losses per unit to identify the suitability of the *nominal-the-best* loss function (equation [3]) for asymmetric cases. All tables include the bootstrap simulation output for three specification settings, *i.e., 1) symmetric specification of the nominal-the-best loss function (n = 100), 2) upper (n = 57), and 3) lower side (n = 43) for asymmetric specification of the nominal-the-best loss function.* To compute the losses and statistics for each side of the asymmetric case the values below and above the target (13.5%) formed an individual dataset.

Table 13 presents the means and variances for the original sample and the bootstrap simulation. The mean value for the bootstrap statistic (N = 10,000) for the symmetric case is the same as for the original sample with 13.94 mm. In contrast the mean values for the bootstrap statistics and original samples for the asymmetric cases are different. For example, the grand-mean value from the bootstrap simulation for the asymmetric lower-side case is 12.50 mm and the mean of the original sample is 12.51 mm (Table 13). A negligible difference of just 0.01 mm. In contrast, the variance values for the bootstrap simulation differ from variances of the original samples for all three cases. Compare the asymmetric upper side case (Table 13), the bootstrap variance is 0.266 mm and variance for the original sample is 0.289 mm. These results suggest that a small bias for the mean and variance values exist.

The cost constants k and the average losses per unit calculated with the bootstrap statistics for all three cases from Table 13 are shown in Table 14. The average loss per unit for the symmetric case is \$10.32 (k = \$5 $(\%^2)^{-1}$), for the asymmetric lower side \$9.61 (k =

Table 13. Comparison of statistics for original sample and bootstrap data for symmetric and asymmetric *nominal-the-best* quality characteristics moisture content (woody residues).

Case	Symmetric		Asymmetric			
			Lower Side		Upper Side	
Original sample size	100		43		57	
Bootstrap Statistic	No	Yes	No	Yes	No	Yes[3]
Mean in [%]	13.939	13.939	12.513	12.504	15.014	15.021
Variance in [%²]	2.082	2.061	0.289	0.266	0.727	0.690

Table 14. Average loss per unit and cost constant *k* for symmetric and asymmetric using bootstraps statistics for *nominal-the-best* quality characteristic.

Case	Symmetric[1]	Asymmetric[2]	
		Lower Side	Upper Side
Original sample size	100	43	57
k [$/%²]	5	8.89	3.2
Average loss per unit [$]	10.32	11.18	9.61

[1] Target = 14 mm; USL = 16mm; LSL = 12mm
[2] Target = 13.5 mm; USL = 16mm; LSL = 12mm

$3.2 $(\%^2)^{-1}$), and for the asymmetric upper side is $11.18 ($k$ = $8.89 $(\%^2)^{-1}$) (Table 14). The average losses per unit were computed with the equations [3] (symmetric case), [15] and [17] (asymmetric cases).

Despite having a smaller cost constant k, less than half, the average loss per unit for the upper side is almost as high as for the lower side. This stems from a two and a half times larger variance for the upper side data (Table 13). The average loss per unit for the symmetric case is mainly driven by the variance of the data around the target. The loss is based on a small difference (0.06mm) of the mean from the target (14%) and a variance of 2.061mm. This could support the statements made in example two, that the *nominal-the-best* quality loss function is more sensitive towards the variance than to shifts in the mean.

Table 15 presents the total losses for all three cases. The first total loss presented (third row in the table) is calculated based on the average loss per unit, *i.e., the average loss per unit from Table 14 times the number of samples of the specific data set.* The second total loss presented (fourth row in the table) is the sum of all losses based on the individual quality characteristic value. The equations [2] (symmetric), [5] and [6] (asymmetric) were used to compute the individual losses. Since Taguchi presented his *nominal-the-best* quality loss function to compute the average loss per unit for symmetric cases, the question was if this equation [3], adjusted, could be used to calculate the average loss per unit for each side for asymmetric cases. Thus, both approaches were compared to investigate the suitability of computing the total loss based on the average loss per unit for asymmetric *nominal-the-best* loss functions. Often the literature suggests for asymmetric cases to compute the loss based on the sum of all individual losses or more complex equations.

The total loss based on the average loss per unit for the symmetric case is $1032, while the total loss for the individual values is $1032.26. The difference (fifth row in the table) is only $0.26 or in other words the difference is 1%. This suggests that equation [2] is a good estimator of the total loss with emphasize on the data variation. The total losses for the lower side are $480.80 to $479.90 with a difference of $0.90. The losses for the upper side are $547.90 to $547.20 with a difference $0.70. These results suggest, that Taguchi's *nominal-the-best* quality loss function (*i.e.,* for more than one unit) is suitable to compute the total loss of the variation for asymmetric settings.

Table 15. Total loss comparison for symmetric and asymmetric *nominal-the-best* quality characteristics.

Case	Symmetric	Asymmetric	
		Lower Side	Upper Side
Original sample size	100	43	57
Total loss[1] based on average loss per unit in [$]	1032.00	480.80	547.90
Total loss[2] based on sum of all individual losses in [$]	1032.26	479.90	547.20
Difference in [$]	0.26	0.90	0.70

[1] Recall equation [2]; $L = k * (\sigma^2+(\bar{y} - m)^2)$
[2] Recall equation [1]; $L(y) = k * (y - m)^2$

CHAPTER FIVE
CONTINUOUS IMPROVEMENT HANDBOOK

This continuous improvement handbook (Appendix C and Attachment File 3) is intended to introduce statistical process control procedures, lean management tools, and Taguchi's quality loss functions to practitioners of the sustainable bio-based industries. This handbook shall function as a useful guide to monitor and reduce material or process variation. The following pages are intended as an introduction to use of the handbook (Figure 29).

The handbook starts with outlining recent economic developments for the sustainable bio-based industries and emphasizes the importance of variation in manufacturing. Variation is crucial since it exists in every component of the production and is an important factor in determining the success of an enterprise, *i.e., differences in material quality, process execution, or even human actions affect the companies' performance.* Therefore, being able to visualize, detect, and quantify variation is vital for the competitiveness of an enterprise. Practitioners are encouraged to apply the developed simulation tool to quantify the variation of the components of their production systems.

Continuous improvement is introduced as a never-ending process and philosophy of little steps towards incremental improvement of the companies' production performance. For a successful application of the continuous improvement philosophy companies of the bio-based industries must fulfill certain requirements first. Continuous improvement is a philosophy or a culture which all entities (*e.g.*, management, workers, etc.) of the enterprise must live, *i.e., the success of continuous improvement is significantly hampered if this requirement is not fulfilled.* Furthermore, continuous improvement requires a great IT-infrastructure for data storage and application of real-time data mining (Young, 2015). The results and inferences drawn from statistical methodologies and tests must be accepted by management and workers (Young, 2015). The following paragraphs give a brief introduction and provide a sequence for the application of key continuous improvement techniques.

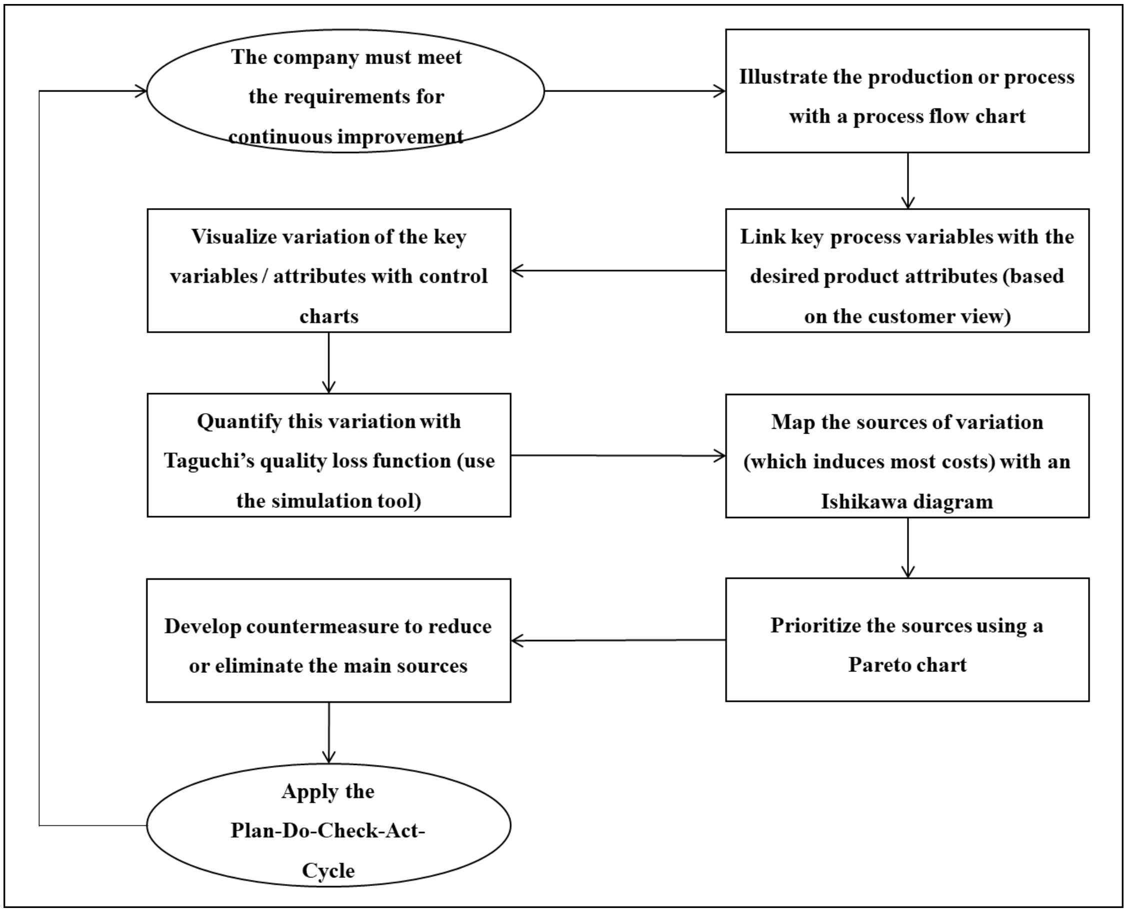


Figure 29. Flow chart on how to use the handbook.

Start continuous improvement by using the **process flow charting** technique to illustrate the logical sequence of all components of the production or process. Standardized symbols are usually used to represent specific actions, *e.g., a rectangular emphasizes one component (process step).* Subsequently, **link key process variables** with by the customer desired **main product attributes** (Young, 2015). Changes in key process variables have direct impact on the product attribute. Linking process variables with product attributes emphasizes the production of products based on the customers' view, *i.e., ask is the customer willing to pay for the product or service.*

The next part of the handbook introduces the reader to **relevant statistics** used to describe data distributions such as the mean, standard deviation, or variance, etc. **Histograms** are used to show data distributions (*e.g.,* normal distribution). For **normal distributions** roughly 99.7% of the data values lie between three standard deviations. The control limits of control charts represent three standard deviation. Shewhart's **control chart** is a tool to visualize **natural-cause** and **special-cause variation**. The control limits distinguish both types of variation based on the three-standard deviation. Common univariate control charts and run rules to detect special cause variation are presented. X-individual and moving range charts are used to provide examples for control charting.

The next chapter discusses the fundamental difference between the traditional quality view and the continuous quality view developed by **Genichi Taguchi**. Traditional quality is defined as conformity to specification, *i.e., all products within specification limits are equally good and cause no loss.* In contrast Taguchi's view on quality is that every product deviating from the target causes loss, *i.e., the further the deviation the higher the loss.* Page 15 of the handbook gives an overview of the three **quality loss functions** provided by Taguchi. The developed simulation model uses these quality loss functions to quantify the variation of components of a series system to identify the component inducing the most costs in the system. Taguchi's quality loss functions are explained with an example on the page 16.

Variation reduction starts by identifying the sources of the visualized and quantified variation. The **Ishikawa diagram** helps to categorize the sources (not symptoms) causing the variation; brainstorm as a team. After mapping the sources apply the **pareto chart** to prioritize the sources. Roughly 80% of the variation can assigned to 20% of the sources. Based on the information given by the pareto chart countermeasures can be developed to eliminate or at least reduce the top source for variation.

Finally, apply the **Plan-Do-Check-Act-Cycle** to implement the ongoing journey of continuous improvement. Continuous improvement is a never-ending process. As an addition the **Theory of Constraints** is introduced to identify and optimize bottlenecks in the production.

CHAPTER SIX

CONCLUSIONS AND RECOMMENDATIONS

This study developed strategies for continuous improvement and improved competitiveness of the sustainable bio-based industries. A simulation tool and a continuous improvement handbook applying 'lean methods', statistical process control, and Taguchi's quality loss functions were developed to support practitioners in their efforts to reduce variation from feedstock supplies.

The simulation model is a great tool to quantify the financial loss induced by variation of key feedstock quality characteristics for the biomass supply chain and its components and to identify the component creating the greatest loss in the system. The handbook is a useful manual for practitioners introducing techniques to analyze, visualize, and quantify variation. Continuous improvement techniques are suitable tools to quantify feedstock variation of the sustainable bio-based industries.

Three empirical examples were used to illustrate the capability of the Excel simulation tool and Taguchi's quality loss functions *nominal-the-best* and *smaller-the-better*. Example number two emphasizes Galton's theory of cumulative variances. Sensitivity analyses were conducted on the simulation outputs.

Example One. Industrial data for the ash content of Switchgrass from the harvest and collection operation were quantified with Taguchi's *smaller-the-better* loss function using the Excel simulation tool. The average loss per unit was found to be $17.37 per dry ton with a cost constant k of $1.25 $(\%^2)^{-1}$. The annual total loss for the assumed biorefinery would be $4.3 million dollars, *i.e., demand of 250,000 dry tons cellulosic biomass times $17.37 per dry ton*. The *smaller-the-better* loss function is more sensitive towards shifts in the mean and changes in the USL than to changes in the variance and customer loss at the upper specification limit. Great cost savings can be achieved through reduced mean values of the quality characteristic. Nevertheless, variation must be reduced first for data with a mean approaching the desired target (*i.e.,* zero). The average loss per unit responds in a quadratic pattern for shifts in the mean and changes in the upper specification limit. In contrast, the average loss per unit showed a linear pattern for changes of the variance and customer loss at upper specification limit. Furthermore, samples with quality characteristic values outside of the set specification limits induce significant higher losses than data within the specification limit.

Example Two. Simulated data for the moisture content of woody residues for an example biomass supply chain were quantified with Taguchi's *nominal-the-best* loss function for different variances based on Galton's theory of *'cumulative variances'*. The average loss per unit for the *nominal-the-best* loss function is influenced more by changes of the variance than to shifts in the mean for the given data output. Generally, both the mean and variance of a quality characteristic impact Taguchi's loss. The highest loss for the supply chain is experienced at *densification* with \$13.23 per dry ton. The annual total loss accumulated for all supply chain operations is \$5.8 million dollars. Furthermore, the loss responds quadratic for shifts in the mean and linear to changes in the variance. Example two showed that applying Galton's theory of *'cumulative variances'* has an influence on the computed Taguchi losses. The average loss per unit using unweighted variances is much higher than for independent components, due to the simple addition of the variances and covariances. For example, the loss for *densification* independently computed is \$13.34 (σ^2 = 2.12%^{2}) compared to the loss for unweighted variances is \$48.30 (σ^2 = 9.13%^{2}). The average losses per unit using Galton's weighted variances indicated much smaller losses for component two, three, and four of the series compared with the losses based on Galton's unweighted variances. Galton's theory can provide a better understanding about the dependencies of the different variances in a series system. However, to justify the application of Galton's theory the simulation must be repeated with real data from a supply chain.

Example Three. Simulated data for the particle size of woody residues for one component with symmetric and asymmetric specification settings was quantified with Taguchi's *nominal-the-best* loss function. For example, the estimated total loss for the upper side of the asymmetric loss function for simulated data is \$480.80 compared to the more precise total loss based on the sum of individual losses \$479.90. Resulting in a neglectable difference of \$0.90. Thus, applying the *nominal-the-best* quality loss function to quantify loss based on variation of a data set is suitable for asymmetric specifications.

The results of this study suggest that using Taguchi's quadratic quality loss functions to be a good fit for computing the loss for feedstock quality characteristics based on variation. Furthermore, the use of Taguchi's quality loss function (*i.e., nominal-the-best* and *smaller-the-better*) emphasized the impact of variation in quality characteristics of cellulosic biomass on the supply chain operations and cost. Variation must be understood as the deviation of the average value from the target and the variability around such average of the quality characteristic. Thus, enterprises should strive to reduce variation of the quality characteristic or process while shifting the mean closer to the desirable target. The simulation tool and handbook will help practitioners in the industry to quantify the individual and total losses for their production system. Furthermore, applying sensitivity analysis will help the industry to understand how variation and Taguchi's quality loss functions impact loss.

A major limitation of this research is the lack of industrial data for each quality characteristic for the various components of the biomass supply chain. Given industrial data, the biomass supply chain and its components could be quantified, using the developed simulation model, based on variation of the key quality characteristics ash content, moisture content, and particle size. This quantification would allow enterprises to identify the quality characteristic inducing most loss for each component, as well as to identify the component which induces the most loss for the whole biomass supply chain. Based on the gained knowledge engineers and managers could apply the continuous improvement techniques presented in this thesis and handbook to develop strategies for variation reduction to achieve great cost savings. Furthermore, a comparison of different types of linear and quadratic loss functions may help practitioners develop their own loss function applicable for their processes.

LIST OF REFERENCES

AL-AOMAR, R. A. 2011. Applying 5S lean technology: An infrastructure for continuous process improvement. *World Academy of Science, Engineering and Technology,* 59**,** 2014 - 2019.

ALEXANDER, M. & KUSLEIKA, D. 2016. *Excel 2016 power programming with vba,* Indianapolis, IN, John Wiley & Sons, Inc.

ANDRÉ, N. & YOUNG, T. M. 2013. Real-time process modeling of particleboard manufacture using variable selection and regression methods ensemble. *European Journal of Wood and Wood Products,* 71**,** 361-370.

ANGUS-HANKIN, C., STOKES, B. & TWADDLE, A. 1995. The transportation of fuel wood from forest to facility. *Biomass and Bioenergy,* 9**,** 191-203.

ANTONY, J. 1997. Experiments in quality. *Manufacturing Engineer,* 76**,** 272-275.

ARGO, A. M., TAN, E. C. D., INMAN, D., LANGHOLTZ, M. H., EATON, L. M., JACOBSON, J. J., WRIGHT, C. T., MUTH, D. J., WU, M. M., CHIU, Y.-W. & GRAHAM, R. L. 2013. Investigation of biochemical biorefinery sizing and environmental sustainability impacts for conventional bale system and advanced uniform biomass logistics designs. *Biofuels, Bioproducts and Biorefining,* 7**,** 282-302.

ASTNER, A. F., YOUNG, T. M. & BOZELL, J. J. 2015. Lignin yield maximization of mixed biorefinery feedstocks by organosolv fractionation using Taguchi Robust Product Design. *Biomass and Bioenergy,* 73**,** 209-216.

AWUDU, I. & ZHANG, J. 2012. Uncertainties and sustainability concepts in biofuel supply chain management: A review. *Renewable and Sustainable Energy Reviews,* 16**,** 1359-1368.

BASSAM, N. E. 1998. *Energy plant species: Their use and impact on enviornment and development.,* London, UK, James & James (Science Publishers).

BEAMON, B. M. 1998. Supply chain design and analysis: Models and methods. *International Journal of Production Economics,* 55**,** 281 - 294.

BHUIYAN, N. & BAGHEL, A. 2005. An overview of continuous improvement: from the past to the present. *Management Decision,* 43**,** 761-771.

BITRA, V. S. P., WOMAC, A. R., CHEVANAN, N., MIU, P. I., IGATHINATHANE, C., SOKHANSANJ, S. & SMITH, D. R. 2009. Direct mechanical energy measures of hammer mill comminution of switchgrass, wheat straw, and corn stover and analysis of their particle size distributions. *Powder Technology,* 193**,** 32-45.

BLACKSTONE, J. H. 2001. Theory of constraints - a status report. *International Journal of Production Research,* 39**,** 1053-1080.

BLS, U. S. B. O. L. S. 2018. *Employment development of the wood product industry* [Online]. [Accessed 07 / 01 / 2018].

CAMBERO, C., SOWLATI, T., MARINESCU, M. & RÖSER, D. 2015. Strategic optimization of forest residues to bioenergy and biofuel supply chain. *International Journal of Energy Research,* 39**,** 439-452.

CANTOR, R. A. & RIZY, C. G. 1991. Biomass energy: Exploring the risks of commercialization in the united states of america. *Bioresource Technology,* 35**,** 13.

CARTY, D. M., YOUNG, T. M., ZARETZKI, R. L., GUESS, F. M. & PETUTSCHNIGG, A. 2015. Predicting and correlating the strenght properties of wood composite process parameters by use of boosted regression tree models. *Forest Products Journal*, 65**,** 365-371.

CASLER, M. D., TOBIAS, C. M., KAEPPLER, S. M., BUELL, C. R., WANG, Z.-Y., CAO, P., SCHMUTZ, J. & RONALD, P. 2011. The Switchgrass genome: tools and strategies. *The Plant Genome Journal,* 4.

CHEN, C.-H. 2004. Determining the optimum process mean of a one-sided specification limit with the linear quality loss function of product. *Journal of Applied Statistics*, 31**,** 693-703.

DAI, J., CUI, H. & GRACE, J. R. 2012. Biomass feeding for thermochemical reactors. *Progress in Energy and Combustion Science,* 38**,** 716-736.

DAOUTIDIS, P., MARVIN, W. A., RANGARAJAN, S. & TORRES, A. I. 2013. Engineering biomass conversion processes: a systems perspective. *AIChE Journal*, 59**,** 3-18.

DARR, M. J. & SHAH, A. 2014. Biomass storage: an update on industrial solutions for baled biomass feedstocks. *Biofuels,* 3**,** 321-332.

DEMING, W. E. 1982. *Quality productivity and competitive position,* Cambridge, MA, Massachusetts Institute of Technology, Center for Advanced Engineering Study.

DEMING, W. E. 1986. *Out of the crisis*, Cambridge, MA, Massachusetts Institute of Technology, Center for Advanced Engineering Study. 507p.

DEMING, W. E. 1993. *The new economics,* Cambridge, MA, Massachusetts Institute of Technology, Center for Advanced Engineering Study.

DOE, U. S. D. O. E. 2018. Alternative fuel price report. Retrieved 7/28/2018 https://www.afdc.energy.gov/fuels/prices.html

DOSHI, J. A., KAMDAR, J. D., JANI, S. Y. & CHAUDHARY, S. J. 2012. Root cause analyiss using ishikawa diagram for reducing radiator rejection. *Engineering Researcha nd Applications (IJERA),* 2.

DOUGLAS, B. 2010. Canada report on bioenergy 2010.

EGGINK, A., PALMER, K., SEVERY, M., CARTER, D. & JACOBSON, A. 2018. Utilization of wet forest biomass as both the feedstock and electricity source for an integrated biochar production system. *Applied Engineering in Agriculture,* 34**,** 125-134.

EISA 2007. Energy independence and security act of 2007. Public Law 110 - 140.

EKŞIOĞLU, S. D., ACHARYA, A., LEIGHTLEY, L. E. & ARORA, S. 2009. Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers & Industrial Engineering,* 57**,** 1342-1352.

EMERY, I. R. & MOSIER, N. S. 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. *Biomass and Bioenergy,* 39**,** 237-246.

EMILIANI, M. L. 2008. Standardized work for executive leadership. *Leadership & Organization Development Journal,* 29**,** 24-46.

EPA 2007. Biomass conversion: Emerging technologies, feedstocks, and products. *In:* DEVELOPMENT, O. O. R. A. (ed.). Washingotn, D.C.: Sustainability Program, Office of Research and Development, U.S. Environmental Protection Agency.

FAO, F. A. A. O. O. T. U. N. 2008. The state of food and agriculture. Rome.

GALIK, C. S., ABT, R. & WU, Y. 2009. Forest biomass supply in the southeastern unitd states - implications for industrial roundwood and bioenergy production. *Journal of Forestry,* 107**,** 69 - 77.

GEORGE, M. L. 2002. *Lean Six Sigma: combining six sigma quality with lean production speed*, McGraw-Hill.

GERMAIN, R., CLAYCOMB, C. & DRÖGE, C. 2008. Supply chain variability, organizational structure, and performance: The moderating effect of demand unpredictability. *Journal of Operations Management,* 26**,** 557-570.

GOLD, S. & SEURING, S. 2011. Supply chain and logistics issues of bio-energy production. *Journal of Cleaner Production,* 19**,** 32-42.

GOLDRATT, E. M. 1990. *What is this thing called theory of constraints and how should it be implemented?,* North River Press.

GRANT, R. M., SHAMI, R. & KRISHNAN, R. 1994. TQM's challenge to management theory and practice. *Sloan Management Review,* 35**,** 25-35.

GUO, M., SONG, W. & BUHAIN, J. 2015. Bioenergy and biofuels: History, status, and perspective. *Renewable and Sustainable Energy Reviews,* 42**,** 712-725.

HAMERNIK, J. 2018. Sources of variation of treated wood. Fachhochschule Salzburg University of Applied Sciences.

HANSEN, E. N. 2010. The role of innovation in the forest products industry. *Journal of Forestry***,** 348 - 353.

HARRY, M. & SCHROEDER, R. 2000. *Six sigma - the breakthrough management strategy revolutionizing the world's top corporations,* New York, NY, Doubleday.

HESS, J. R., WRIGHT, C. T. & KENNEY, K. L. 2007. Cellulosic biomass feedstocks and logistics for ethanol production. *Biofuels, Bioproducts and Biorefining,* 1**,** 181-190.

HESS, J. R., WRIGHT, C. T., KENNEY, K. L. & SEARCY, E. M. 2009. Uniform-format solid feedstock supply system: A commodity-scale design to produce an infrastructure-compatible bulk solid from lignocellulosic biomass. *Idaho National Laboratory,* INL/EXT-09-15423.

HESTERBERG, T. C. 1998. Simulation and bootstrapping fo teaching statistics.

HIMMEL, M., TUCKER, M., BAKER, J., RIVARD, C., OH, K. & GROHMANN, K. 1985. Comminution of biomass: hammer and kinfe mills. *Proceedings of the 7th Symposium on Biotechnology for Fuels and Chemicals.* Gatlinburg, Tennessee: John Wiley.

HINES, P. & RICH, N. 1997. The seven value stream mapping tools. *International Journal of Operations & Production Management,* 17**,** 46-64.

HOEKMAN, S. K. 2009. Biofuels in the U.S. – challenges and opportunities. *Renewable Energy,* 34**,** 14-22.

HOWARD, J. L. & JONES, K. C. 2016. U.S. Timber production, trade, consumption, and price statistics, 1965-2013. Madison, WI: U.S. Department of Agriculture, Forest Service, Forest Products Laboratory.

IEA BIOENERGY 2007. Potential contribution of bioenergy to the world's future energy demand.

IGATHINATHANE, C., WOMAC, A. R., SOKHANSANJ, S. & NARAYAN, S. 2008. Knife grid size reduction to pre-process packed beds of high- and low-moisture switchgrass. *Bioresour Technol,* 99**,** 2254-64.

INSTITUTE, M. R. 1999. Emission factor documentation for AP-42 section 10.8 - wood preserving - final report.

ISHIKAWA, K. 1986. *Guide to quality control,* Quality Resources.

ISU 2018. Historical ethanol operating margins. Iowa State University. Retrieved 7/28/2018, from Iowa State University https://www.card.iastate.edu/research/biorenewables/tools/hist_eth_gm.aspx

JACOBSON, J. J., RONI, M. S., CAFFERTY, K. G., KENNEY, K., SEARCY, E. & HANSEN, J. 2014. Feedstock supply system design and analysis. Idaho National Laboratory.

JURAN, J. M. 1989. *Juran on leadership for quality: an executive handbook,* Free Press.

JURAN, J. M. & GRYNA, F. M. 1951. *Juran's quality control handbook,* New York, NY, McGraw-Hill.

JURAN, J. M. & GRYNA, F. M. 1993. *Quality planning and analysis,* New York, NY, Mcgraw-Hill.

KEEFE, R., ANDERSON, N., HOGLAND, J. & MUHLENFELD, K. 2014. Woody biomass logistics. *Cellulosic Energy Cropping Systems.*

KEEGAN, C. E., SORENSON, C. B., MORGAN, T. A., HAYES, S. W. & DANIELS, J. M. 2011. Impact of the great recession and housing collapse on the forest products industry in the western united states. *Forest Products Journal,* 61**,** 625-634.

KENNEY, K. L., SMITH, W. A., GRESHAM, G. L. & WESTOVER, T. L. 2013. Understanding biomass feedstock variability. *Biofuels,* 4**,** 111-127.

KIM, M. W. & LIAO, W. M. 1994. Estimating hidden quality costs with quality loss functions. *American Accounting Association - Accounting Horizons,* 8**,** 8 - 18.

LACEY, J. A., EMERSON, R. M., WESTOVER, T. L. & THOMPSON, D. N. 2016. Ash reduction strategies in corn stover facilitated by anatomical and isze fractionation. *In:* LABORATORY, I. N. (ed.).

LAM, P. S., SOKHANSANJ, S., BI, X., LIM, C. J., NAIMI, L. J., HOQUE, M., MANI, S., WOMAC, A. R., YE, X. P. & NARAYAN, S. 2008. Bulk density of wet and dry wheat straw and switchgrass particles. *Applied Engineering in Agriculture,* 24**,** 351-358.

LAMERS, P., RONI, M. S., TUMULURU, J. S., JACOBSON, J. J., CAFFERTY, K. G., HANSEN, J. K., KENNEY, K., TEYMOURI, F. & BALS, B. 2015. Techno-economic analysis of decentralized biomass processing depots. *Bioresour Technol,* 194**,** 205-13.

LEAVENGOOD, S. & REEB, J. 2002. Part 3: Pareto analysis and check sheets. *Statistical process control.* Corvallis, Or.: Extension Service, Oregon State University.

LI, C., ASTON, J. E., LACEY, J. A., THOMPSON, V. S. & THOMPSON, D. N. 2016. Impact of feedstock quality and variation on biochemical and thermochemical conversion. *Renewable and Sustainable Energy Reviews,* 65**,** 525-536.

LI, M.-H. C. 2002. Unbalanced tolerance design and manufacturing setting with asymmetrical linear loss function. *The International Journal of Advanced Manufacturing Technology,* 20**,** 334-340.

LIAO, C.-N. 2010. Supplier selection project using an integrated delphi, ahp and taguchi loss function. *ProbStat Forum,* 3**,** 118-134.

LIN, T., RODRÍGUEZ, L. F., DAVIS, S., KHANNA, M., SHASTRI, Y., GRIFT, T., LONG, S. & TING, K. C. 2016. Biomass feedstock preprocessing and long-distance transportation logistics. *GCB Bioenergy,* 8**,** 160-170.

LIN, T., RODRÍGUEZ, L. F., SHASTRI, Y. N., HANSEN, A. C. & TING, K. C. 2013. GIS-enabled biomass-ethanol supply chain optimization: model development and Miscanthus application. *Biofuels, Bioproducts and Biorefining,* 7**,** 314-333.

LOFTHOUSE, T. 1999. The Taguchi loss function. *Work Study,* 48**,** 218-223.

MA, Z., WOOD, C. W. & BRANSBY, D. I. 2000. Impacts of soil managment on root characteristics of switchgrass. *Biomass and Bioenergy,* 18**,** 8.

MABEE, W. E., MCFARLANE, P. N. & SADDLER, J. N. 2011. Biomass availability for lignocellulosic ethanol production. *Biomass and Bioenergy,* 35**,** 4519-4529.

MACROTRENDS LLC 2018a. Crude oil prices - 70 year historical chart. Retrieved 7/26/2018 https://www.macrotrends.net/1369/crude-oil-price-history-chart

MACROTRENDS LLC 2018b. Housing starts historical chart. Retrieved 7/26/2018 https://www.macrotrends.net/1314/housing-starts-historical-chart

MANESS, T. C., KOZAK, R. A. & STAUDHAMMER, C. 2003. Applying real-time statistical process control to manufacturing processes exhibiting between and within part size variability in the wood products industry. *Quality Engineering,* 16**,** 113-125.

MARINO, D., AMARASINGHE, K., ANDERSON, M., YANCEY, N., NGUYEN, Q., KENNEY, K. & MANIC, M. 2017. Data driven decision support for reliable biomass feedstock preprocessing. *Resilience Week (RWS)***,** 97 - 102.

MCKENDRY, P. 2002. Energy production from biomass (part 1): overview of biomass. *Bioresource Technology,* 83**,** 37-46.

MCLAUGHLIN, S. B. & ADAMS KSZOS, L. 2005. Development of switchgrass (Panicum virgatum) as a bioenergy feedstock in the United States. *Biomass and Bioenergy,* 28**,** 515-535.

MCLAUGHLIN, S. B., DE LA TORRE UGARTE, D. G., GARTEN, J., C.T., LYND, L. R., SANDERSON, M. A., TOLBERT, V. R. & WOLF, D. D. 2002. High-value renewable energy form prairie grasses. *Environmental Science & Technology,* 36.

MEUNIER-GODDIK, L., BOTHWELL, M., SANGSEETHONG, K., PIYACHOMKWAN, K., CHUNG, Y.-C., THAMMASOUK, K., TANJO, D. & PENNER, M. H. 1999. Physicochemical properties of pretreated poplar feedstocks during simultaneous saccharification and fermentation. *Enzyme and Microbial Technology,* 24**,** 667-674.

MIAO, Z., GRIFT, T. E., HANSEN, A. C. & TING, K. C. 2011. Energy requirement for comminution of biomass in relation to particle physical properties. *Industrial Crops and Products,* 33**,** 504-513.

MITCHELL, R. & SCHMER, M. 2012. Switchgrass harvest and storage. *Switchgrass - A Valuable Biomass Crop for Energy.* London: Springer London Heidelberg New York Dordrecht.

MONTGOMERY, D. C. 2012. *Design and analysis of experiments, 8th edition*, John Wiley & Sons, Inc.

MORIANA, R., VILAPLANA, F. & EK, M. 2015. Forest residues as renewable resources for bio-based polymeric materials and bioenergy: chemical composition, structure and thermal properties. *Cellulose,* 22**,** 3409-3423.

MOSER, L. E. & VOGEL, K. P. 1995. Switchgrass, bigbluestem and indiangrass. *In:* BARNES, R. F., MILLER, D. A. & NELSON, C. J. (eds.) *Forages, an introduction to grassland agriculture.* 1 ed. Ames: Iowa State University Press

NELSON, L. S. 1984. Technical aids. *Journal of Quality Technology,* 16**,** 238-239.

NETTLES, J., BIRKS, P., SUCRE, E. & BILBY, R. 2015. Sustainable production of bioenergy feedstock from the industrial forest: potential and challenges of operational scale implementation. *Current Sustainable/Renewable Energy Reports,* 2**,** 121-127.

NICHOLLS, D. & BUMGARDNER, M. 2018. Challenges and opportunities for north american hardwood manufacturers to adopt customization strategies in an era of increased competition. *Forests,* 9.

OHNO, T. 1988. *Toyota production system: beyond large-scale production,* Cambrigde, MA, Productivity Press.

PARRISH, D. J., CASLER, M. D. & MONTI, A. 2012. The evolution of switchgrass as an energy crop. *In:* MONTI, A. (ed.) *Switchgrass a valuable biomass crop for energy.* London: Springer-Verlag.

PERLACK, R. D., WRIGHT, L. L., TURHOLLOW, A. F. & GRAHAM, R. L. 2005. Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. *In:* AGRICULTURE, U. S. D. O. (ed.).

PLATZER, M. 2016. *"A simulation model of the "bio-depot" concept in the context of components of variance and the "taguchi loss function".pdf".* Master's Thesis, University of Tennessee.

POTTEL, H. 2015. Bootstrapping, permutation tests and simulation in excel, https://www.researchgate.net/profile/Hans_Pottel/publication/278684962_Resampling_methods_in_Excel/links/5583de7308ae89172b860592/Resampling-methods-in-Excel.pdf.

RADNOR, Z. & LESEURE, M. 2010. Transferring lean into government. *Journal of Manufacturing Technology Management,* 21**,** 411-428.

RAND, G. K. 2000. Critical chain: the theory of constraints applied to project management. *International Journal of Project Management,* 18**,** 173-177.

RAY, A. E., LI, C., THOMPSON, V. S., DAUBARAS, D. L., NAGLE, N. J. & HARTLEY, D. S. 2017. Biomass blending and densification: impacts on feedstock supply and biochemical conversion performance. *Biomass Volume Estimation and Valorization for Energy.*

RENTIZELAS, A. A., TOLIS, A. J. & TATSIOPOULOS, I. P. 2009. Logistics issues of biomass: the storage problem and the multi-biomass supply chain. *Renewable and Sustainable Energy Reviews,* 13**,** 887-894.

RIEGLER, M., ANDRÉ, N., GRONALT, M. & YOUNG, T. M. 2015. Real-time dynamic simulation of continuous bulk material flow to improve the statistical modelling of final product strength properties. *International Journal of Production Research,* 53**,** 6629-6636.

ROBERTS, B. J. 2014. Forest residues. *https://www.nrel.gov/gis/images/biomass_2014/National-Biomass-Forest-2014-01.jpg*. National Renewable Energy Laboratories.

ROTHER, M. & SHOOK, J. 1999. *Learning to see: value stream mapping to add value and eliminate muda,* Cambridge, MA, The Lean Enterprise Institute, Inc.

SALIM, R. & JOHANSSON, J. 2016. The influence of raw material on the wood product manufacturing. *Procedia CIRP*, 57**,** 764-768.

SAUERS, D. G. 1999. Using the Taguchi loss function to reduce common-cause variation. *Quality Engineering,* 12**,** 245-252.

SCAVARDA, A. J., BOUZDINE-CHAMEEVA, T., MEYER GOLDSTEIN, S., HAYS, J. M. & HILL, A. V. A review of the causal mapping practice an dresearch literature. Second World Conference on POM and 15th Annual POM Conference, 2004 Cancun, Mexico.

SCHELL, D. J. & HARWOOD, C. 1994. Milling of lignocellulosic biomass: results of pilot-scale testing. *Applied Biochemistry and Biotechnology,* 45**,** 159-168.

SCHNEPF, R. 2011. Cellulosic ethanol: feedstocks, conversion technologies, economics, and policy options. *Asia Pacific Journal of Life Sciences,* vol. 5, no. 3**,** 171-243.

SEARCY, E., WRIGHT, C. T., MUTH, D. J., HESS, J. R. & JACOBSON, J. J. 2010. The limits of a 50 mile collection radius for biomass supply to a biorefinery, draft report. *Idaho National Laboratory, Idaho Falls, ID*.

SHARMA, B., INGALLS, R. G., JONES, C. L. & KHANCHI, A. 2013. Biomass supply chain design and analysis: basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews,* 24**,** 608-627.

SHEWHART, W. A. 1931. *Economic control of quality of manufactured product,* New York, D. Van Nostrand Company. pp. 501 p.

SHEWHART, W. A. 1939. *Statistical method from the viewpoint of quality control,* New York, Dover Publications, Inc.

SHINNERS, K. J., BENNET, R. G. & HOFFMAN, D. S. 2012. Single- and two-pass corn grain and stover harvesting. *American Society of Agricultural and Biological Engineers,* 55**,** 341-350.

SIMS, R. E., MABEE, W., SADDLER, J. N. & TAYLOR, M. 2010. An overview of second generation biofuel technologies. *Bioresour Technol,* 101**,** 1570-80.

SINGH, J. & SINGH, H. 2015. Continuous improvement philosophy – literature review and directions. *Benchmarking: An International Journal,* 22**,** 75-119.

SMITH, W. B., MILES, P. D., PERRY, C. H. & PUGH, S. A. 2009. Forest resources of the United States, 2007. Washington, DC: U.S. Department of Agriculture, Forest Service.

SOFUOGLU, S. D. & KURTOGLU, A. 2012. A survey for determination of wastage rates at massive wood materials processing. *Wood Research,* 57**,** 297-308.

SOKHANSANJ, S. & TURHOLLOW, A. F. 2004. Biomass densification - cubin operations and costs for corn stover. *Applied Engineering in Agriculture,* 20**,** 495 - 499.

SRINIVASAN, M. M. 2011. *Building lean spply chains with the theory of constraints*, McGraw-Hill Education.

STEINER, S., ZENG, Y., YOUNG, T. M., EDWARDS, D. J., GUESS, F. M. & CHEN, C.-H. 2017. A study of missing data imputation in predictive modeling of a wood-composite manufacturing process. *Journal of Quality Technology,* 48**,** 284-296.

STELTE, W., SANADI, A. R., SHANG, L., HOLM, J. K., AHRENFELDT, J. & HENRIKSEN, U. B. 2012. Recent developments in biomass pelletization - a review. *BioResources,* 7**,** 4451-4490.

STIGLER, S. M. 2010. Darwin, Galton and the statistical enlightenment. *Journal of the Royal Statistical Society, Series A,* 173**,** 469-482.

SUNDAR, R., BALAJI, A. N. & KUMAR, R. M. S. 2014. A review on lean manufacturing implementation techniques. *Procedia Engineering,* 97**,** 1875-1885.

TAGUCHI, G. 1993. *Taguchi on robust technology development,* New York, NY, The American Society of Mechanical Engineers. American Society of Mechangical Engineers (ASME) Press.

TAGUCHI, G., CHOWDHURY, S. & WU, Y. 2004. *Taguchi's quality engineering handbook,* Hoboken, New Jersey, John Wiley & Sons, Inc.

TAGUCHI, G., ELSAYED, E. A. & HSIANG, T. C. 1988. *Quality engineering in production systems*, McGraw-Hill.

TAO, G., LESTANDER, T. A., GELADI, P. & XIONG, S. 2012. Biomass properties in association with plant species and assortments I: A synthesis based on literature data of energy properties. *Renewable and Sustainable Energy Reviews,* 16**,** 3481-3506.

TEERAVARAPRUG, J. 2008. Outsourcing and vendor selection model based on taguchi loss function. *Songklanakarin Journal of Science and Technology,* 30**,** 523 - 530.

THOMPSON, D. N., CAMPBELL, T., BALS, B., RUNGE, T., TEYMOURI, F. & OVARD, L. P. 2014. Chemical preconversion: application of low-severity pretreatment chemistries for commoditization of lignocellulosic feedstock. *Biofuels,* 4**,** 323-340.

TSUTSUI, W. M. 1996. W. Edwards Deming and the origins of quality control in Japan. *Journal of Japanese Studies,* 22**,** 295-325.

TUMULURU, J. S., SEARCY, E., KENNEY, K. L., SMITH, W. A., GRESHAM, G. L. & YANCEY, N. A. 2016. *Impact of feedstock supply systems unit operations on feedstock cost and quality for bioenergy applications,* New York, Nova Science Publishers, Inc. .

TUMULURU, J. S., TABIL, L. G., SONG, Y., IROBA, K. L. & MEDA, V. 2014. Grinding energy and physical properties of chopped and hammer-milled barley, wheat, oat, and canola straws. *Biomass and Bioenergy,* 60**,** 58-67.

TUMULURU, J. S., WRIGHT, C. T., HESS, J. R. & KENNEY, K. L. 2011. A review of biomass densification systems to develop uniform feedstock commodities for bioenergy application. *Biofuels, Bioproducts and Biorefining,* 5**,** 683-707.

TYNER, W. E. 2010. What drives changes in commodity prices? Is it biofuels? *Biofuels,* 1**,** 535-537.

ULUTAS, B. 2011. An application of SMED methodology. *World Academy of Science, Engineering and Technology,* 5.

US DEPARTMENT OF ENERGY 2014. Multi-year program plan. US Dpeartment of Energy, Washington, DC, USA.

VERSCHURREN, G. 2014. Bootstrapping with Excel and VBA. Retrieved from https://www.youtube.com/watch?v=BaL14-C8SeI&t=478s (7/25/2018)

VIDAL, B. C., JR., DIEN, B. S., TING, K. C. & SINGH, V. 2011. Influence of feedstock particle size on lignocellulose conversion--a review. *Appl Biochem Biotechnol,* 164**,** 1405-21.

VOGEL, K. P., SARATH, G., SAATHOFF, A. J. & MITCHELL, R. B. 2010. Chapter 17. Switchgrass. *Energy Crops.*

WARNER, E., MORIARTY, K., LEWIS, J., MILBRANDT, A. & SCHWAB, A. 2017. 2015 bioenergy market report. Golden, Colorado: U.S. Department of Energy, Bioenergy Technologies Office.

WEISS, N. D., FARMER, J. D. & SCHELL, D. J. 2010. Impact of corn stover composition on hemicellulose conversion during dilute acid pretreatment and enzymatic cellulose digestibility of the pretreated solids. *Bioresour Technol,* 101**,** 674-8.

WESTERN ELECTRIC COMPANY 1956. *Statistical quality control handbook,* Indianapolis, Indiana, Western Electric Co.

WHEELER, D. J. & CHAMBERS, D. S. 1992. *Understanding statistical process control* Knoxville, TN, SPC Press.

WILKINSON, L. 2006. Revising the pareto chart. *The American Statistician,* 60**,** 332-334.

WILLIAMS, C. L., WESTOVER, T. L., EMERSON, R. M., TUMULURU, J. S. & LI, C. 2015. Sources of biomass feedstock variability and the potential impact on biofuels production. *BioEnergy Research,* 9**,** 1-14.

WILSON, L. 2010. *How to implement lean manufacturing,* McGraw-Hill.

WOMACK, J. P. & JONES, D. T. 1996. *Lean thinking: banish waste and create wealth in your corporation,* Simon & Schuster.

YANCEY, N., TUMULURU, J. S. & WRIGHT, C. T. 2013. Grinding and densification studies on raw and formulated woody and herbaceous biomass feedstocks. *Journal of Biobased Materials and Bioenergy,* 7**,** 549-558.

YORK, D., JIN, K., SONG, Q. & LI, H. Practical root cause analysis using cause mapping. International MultiConference of Engineers and Computer Scientists 2014, 2014 Hong Kong.

YOU, F., TAO, L., GRAZIANO, D. J. & SNYDER, S. W. 2012. Optimal design of sustainable cellulosic biofuel supply chains: Multiobjective optimization coupled with life cycle assessment and input-output analysis. *AIChE Journal,* 58**,** 1157-1180.

YOUNG, T. M. 2015. Process optimization by means of statistics. *In:* WEIGL, M. (ed.). Vienna, Austria: Holzforschung Austria.

YOUNG, T. M., ANDRÉ, N. & OTJEN, J. 2015a. Quantifying the natural variation of formaldehyde emissions for wood composite panels. *Forest Products Journal,* 65**,** 82-84.

YOUNG, T. M., BOND, B. H. & WIEDENBECK, J. 2007. Implementation of a real-time statistical process control system in hardwood sawmills. *Forest Products Journal,* 57**,** 54-62.

YOUNG, T. M., CLAPP, N. E., GUESS, F. M. & CHEN, C.-H. 2014. Predicting key reliability response with limited response data. *Quality Engineering,* 26**,** 223-232.

YOUNG, T. M., LEÓN, R. V., CHEN, C.-H., CHEN, W., GUESS, F. M. & EDWARDS, D. J. 2015b. Robustly estimating lower percentiles when observations are costly. *Quality Engineering,* 27**,** 361-373.

YOUNG, T. M. & WINISTOFER, P. M. 1999. SPC: statistical process control and the forest products industry. *Forest Products Journal,* 49**,** 10.

YUE, D., YOU, F. & SNYDER, S. W. 2013. Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering,* 66**,** 36-56.

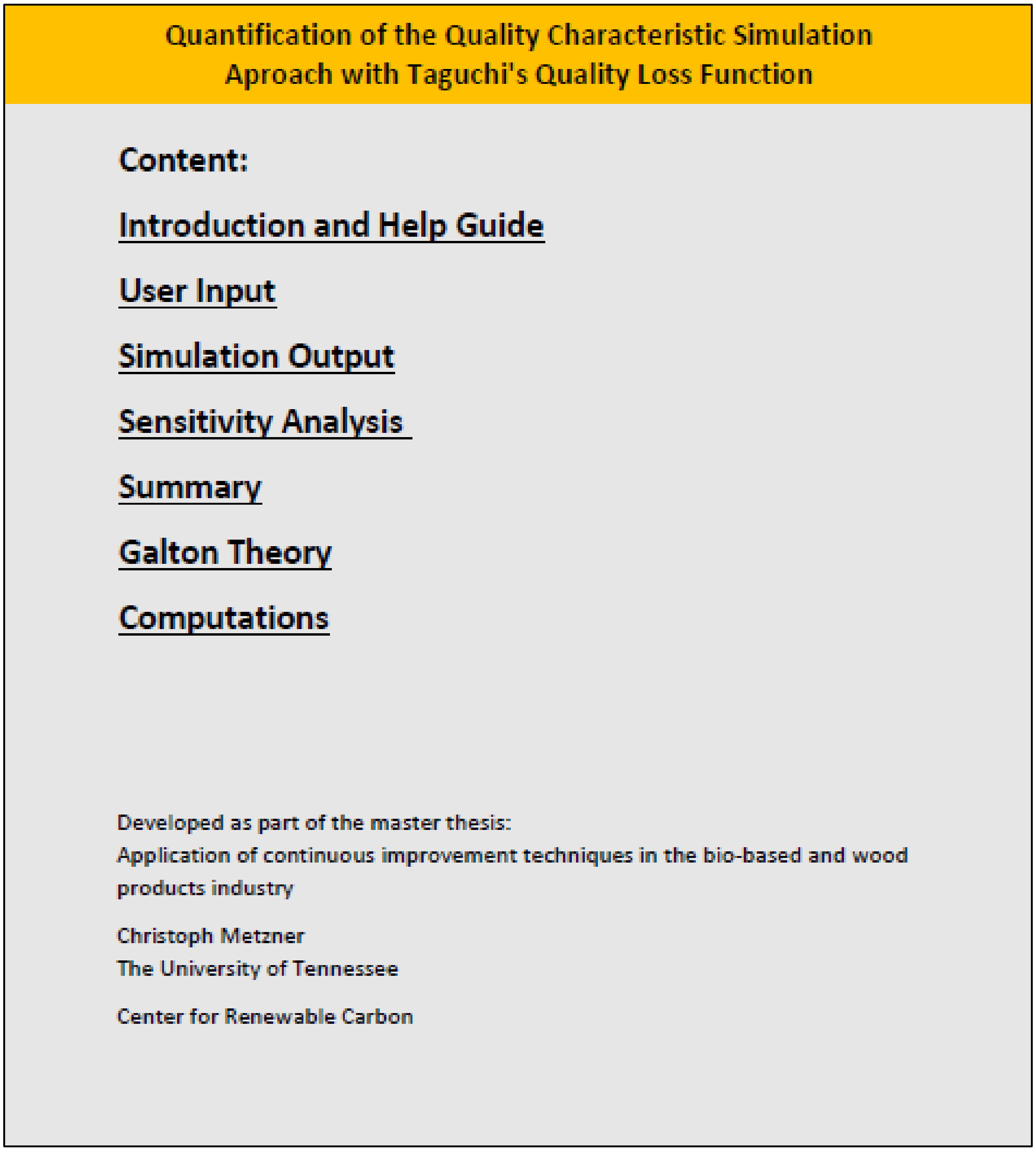
ZANDI ATASHBAR, N., LABADIE, N. & PRINS, C. 2017. Modelling and optimisation of biomass supply chains: a review. *International Journal of Production Research***,** 1-25.

ZEGADA-LIZARAZU, W., WULLSCHLEGER, S. D., NAIR, S. S. & MONTI, A. 2012. Crop physiology. *In:* MONTI, A. (ed.) *Switchgrass a valuable biomass crop for energy.* London: Springer-Verlag.

ZHU, Q. Goal trees and fault trees for root cause analysis. 2008 IEEE International Conference on Software Maintenance, 2008 Beijing, China.

APPENDICES

Appendix A



Quantification of the Quality Characteristic Simulation
Aproach with Taguchi's Quality Loss Function

Content:

Introduction and Help Guide

User Input

Simulation Output

Sensitivity Analysis

Summary

Galton Theory

Computations

Developed as part of the master thesis:
Application of continuous improvement techniques in the bio-based and wood products industry

Christoph Metzner
The University of Tennessee

Center for Renewable Carbon

Figure 30. Excel simulation tool; spreadsheet 1 – content.

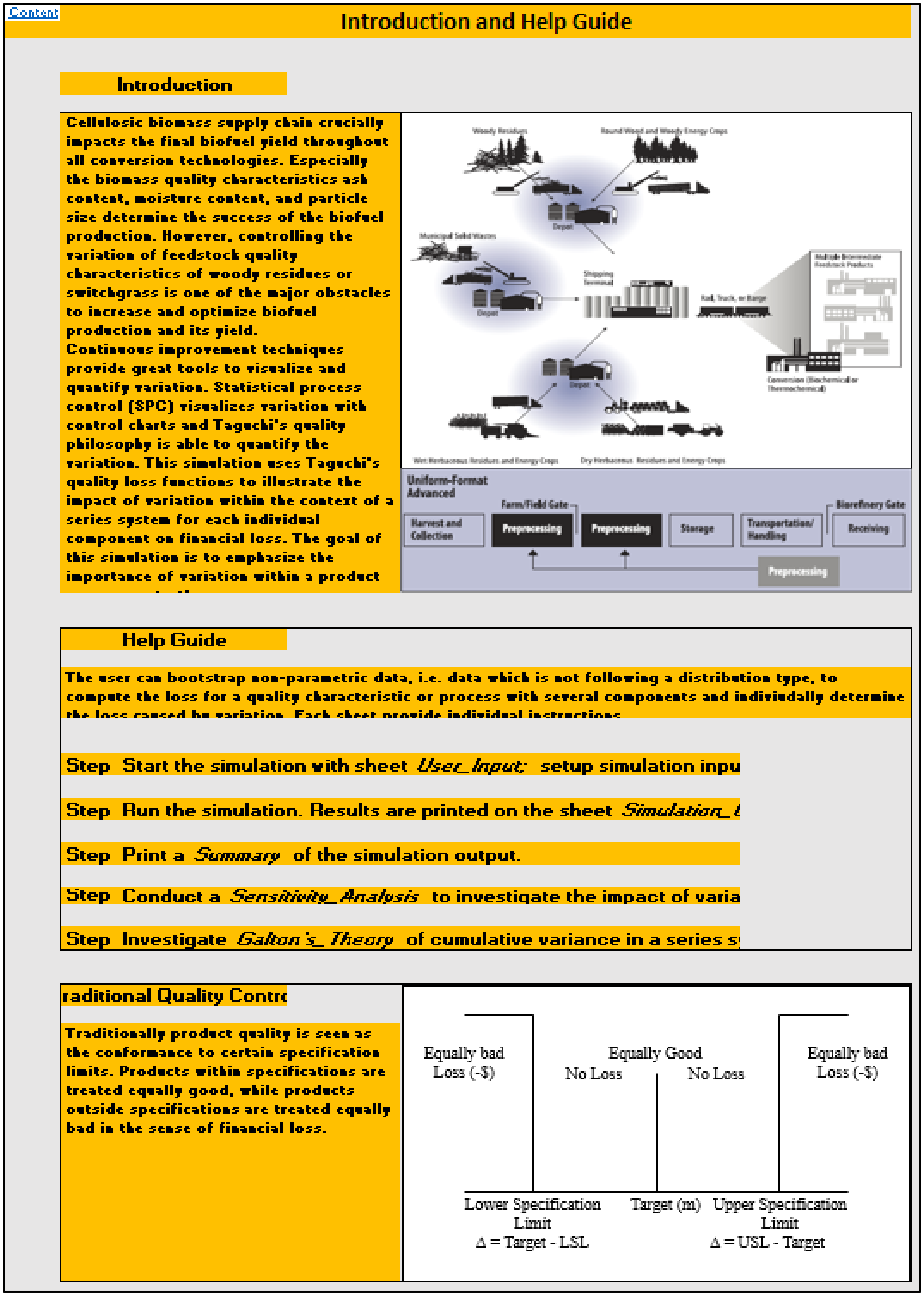
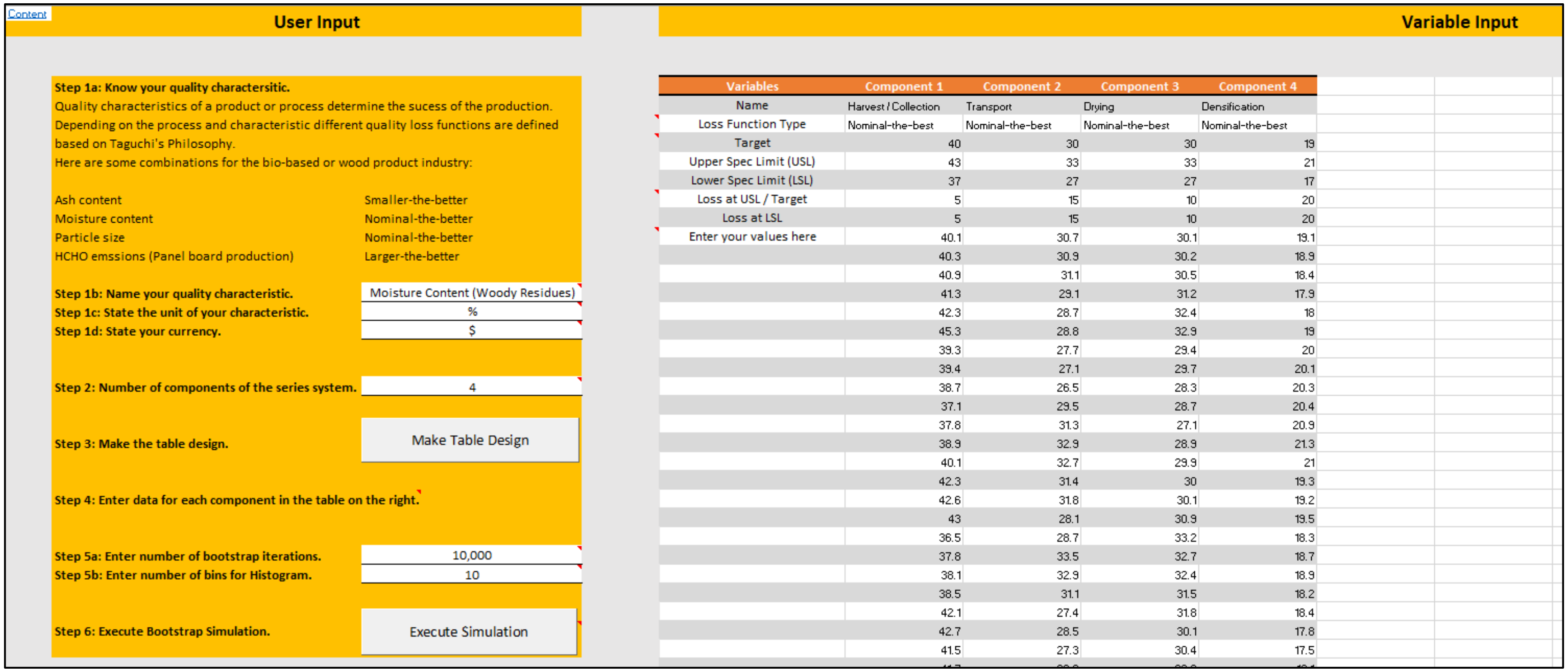


Figure 31. Excel simulation tool; spreadsheet 2 – introduction and help guide.



Content

User Input

Step 1a: Know your quality charactersitic.
Quality characteristics of a product or process determine the sucess of the production.
Depending on the process and characteristic different quality loss functions are defined based on Taguchi's Philosophy.
Here are some combinations for the bio-based or wood product industry:

Ash content	Smaller-the-better
Moisture content	Nominal-the-better
Particle size	Nominal-the-better
HCHO emssions (Panel board production)	Larger-the-better

Step 1b: Name your quality characteristic.	Moisture Content (Woody Residues)
Step 1c: State the unit of your characteristic.	%
Step 1d: State your currency.	$
Step 2: Number of components of the series system.	4
Step 3: Make the table design.	Make Table Design
Step 4: Enter data for each component in the table on the right.	
Step 5a: Enter number of bootstrap iterations.	10,000
Step 5b: Enter number of bins for Histogram.	10
Step 6: Execute Bootstrap Simulation.	Execute Simulation

Variable Input

Variables	Component 1	Component 2	Component 3	Component 4
Name	Harvest / Collection	Transport	Drying	Densification
Loss Function Type	Nominal-the-best	Nominal-the-best	Nominal-the-best	Nominal-the-best
Target	40	30	30	19
Upper Spec Limit (USL)	43	33	33	21
Lower Spec Limit (LSL)	37	27	27	17
Loss at USL / Target	5	15	10	20
Loss at LSL	5	15	10	20
Enter your values here	40.1	30.7	30.1	19.1
	40.3	30.9	30.2	18.9
	40.9	31.1	30.5	18.4
	41.3	29.1	31.2	17.9
	42.3	28.7	32.4	18
	45.3	28.8	32.9	19
	39.3	27.7	29.4	20
	39.4	27.1	29.7	20.1
	38.7	26.5	28.3	20.3
	37.1	29.5	28.7	20.4
	37.8	31.3	27.1	20.9
	38.9	32.9	28.9	21.3
	40.1	32.7	29.9	21
	42.3	31.4	30	19.3
	42.6	31.8	30.1	19.2
	43	28.1	30.9	19.5
	36.5	28.7	33.2	18.3
	37.8	33.5	32.7	18.7
	38.1	32.9	32.4	18.9
	38.5	31.1	31.5	18.2
	42.1	27.4	31.8	18.4
	42.7	28.5	30.1	17.8
	41.5	27.3	30.4	17.5

Figure 32. Excel simulation tool; spreadsheet 3 – user input.

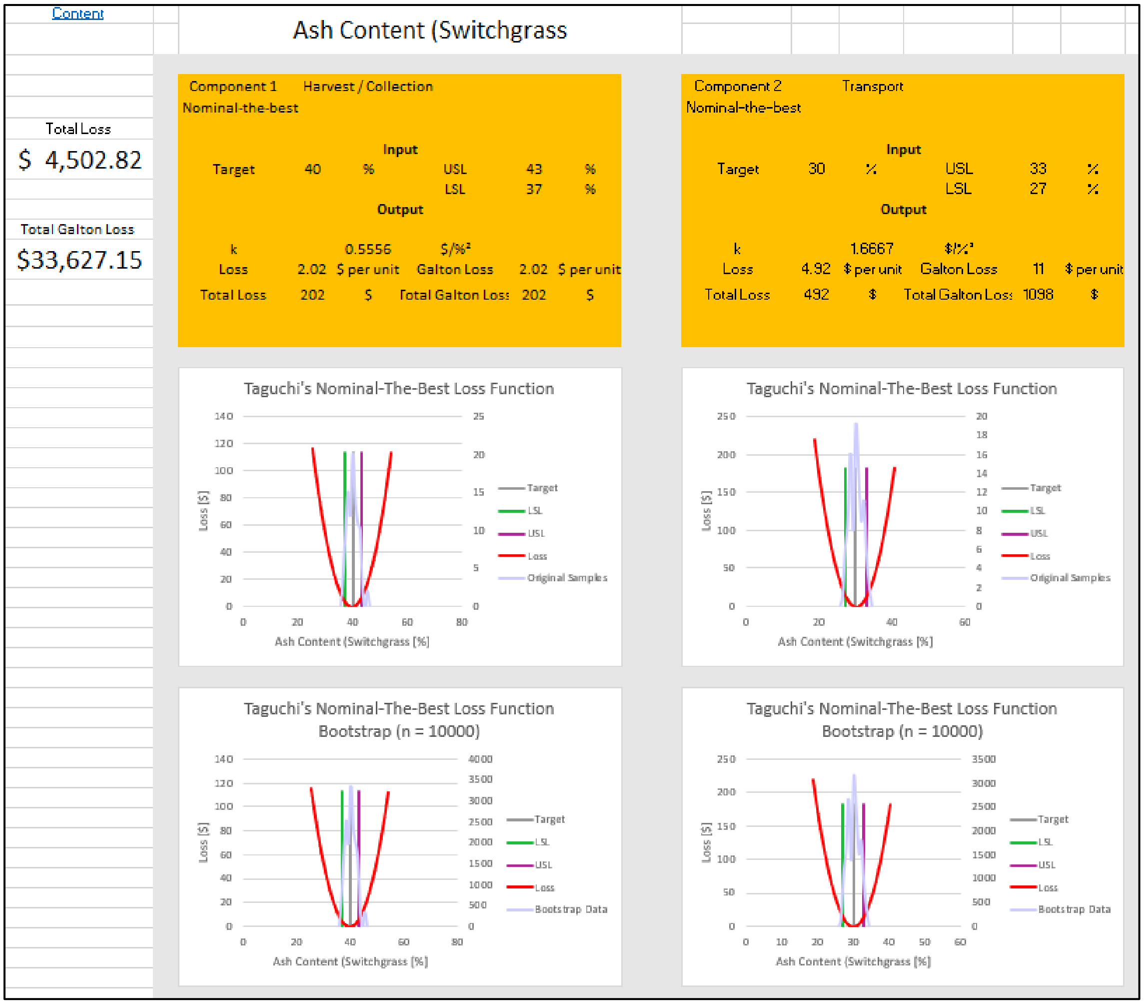


Figure 33. Excel simulation tool; spreadsheet 4 – simulation output – average loss and loss function for component 1 and 2.

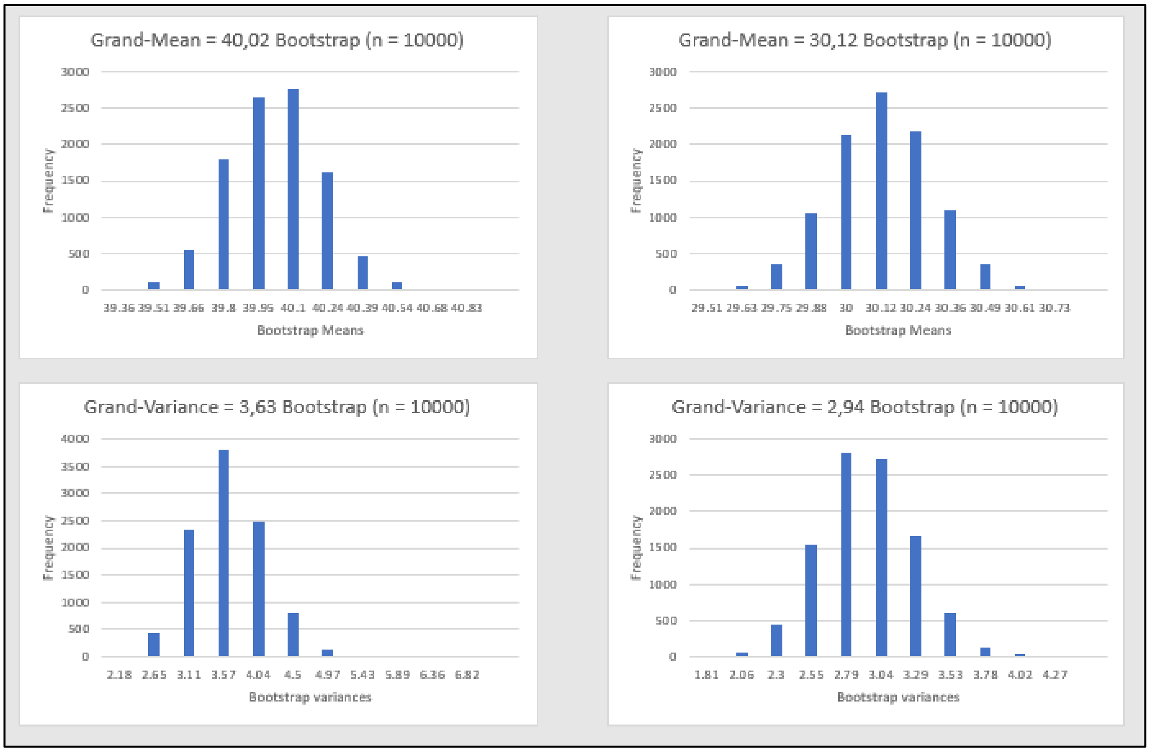


Figure 34. Excel simulation tool; spreadsheet 4 – simulation output (2) – bootstrap statistic distributions for component 1 and 2.

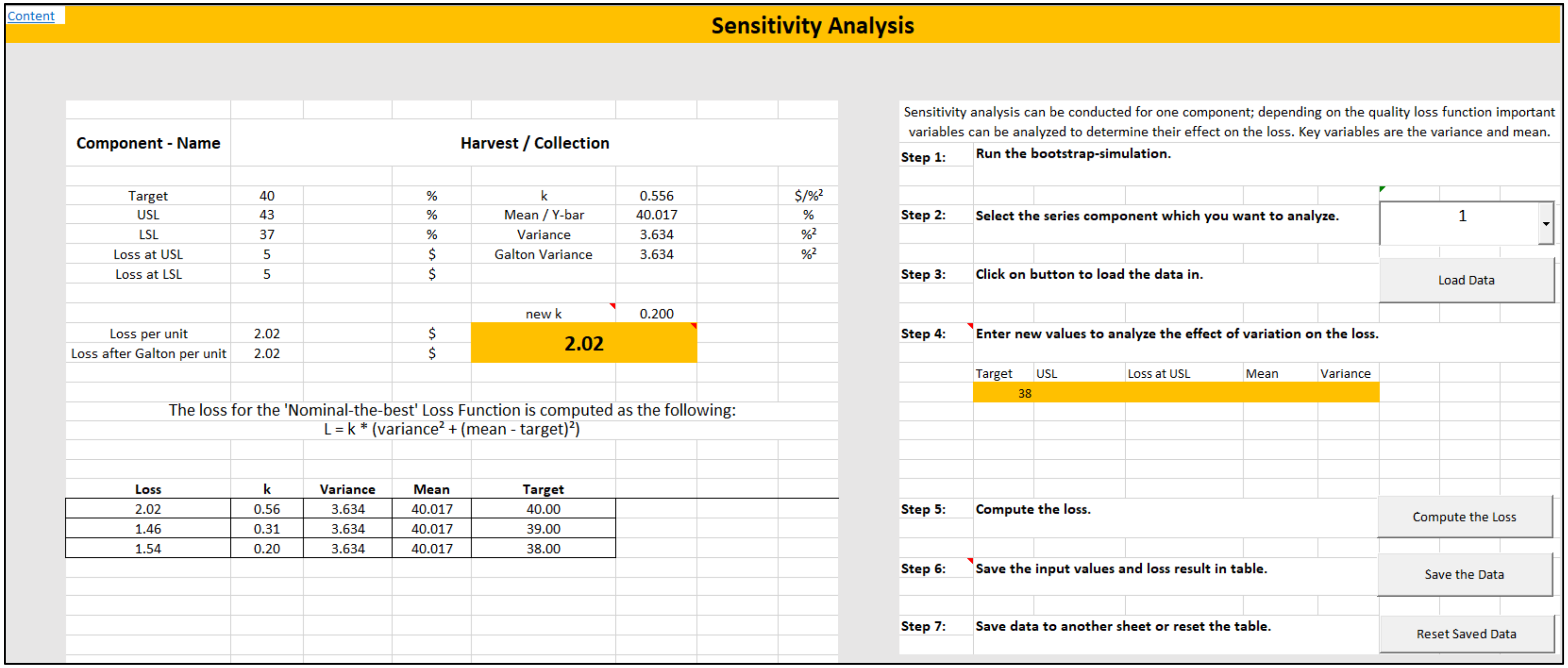


Figure 35. Excel simulation tool; spreadsheet 5 – sensitivity analysis - component 1.

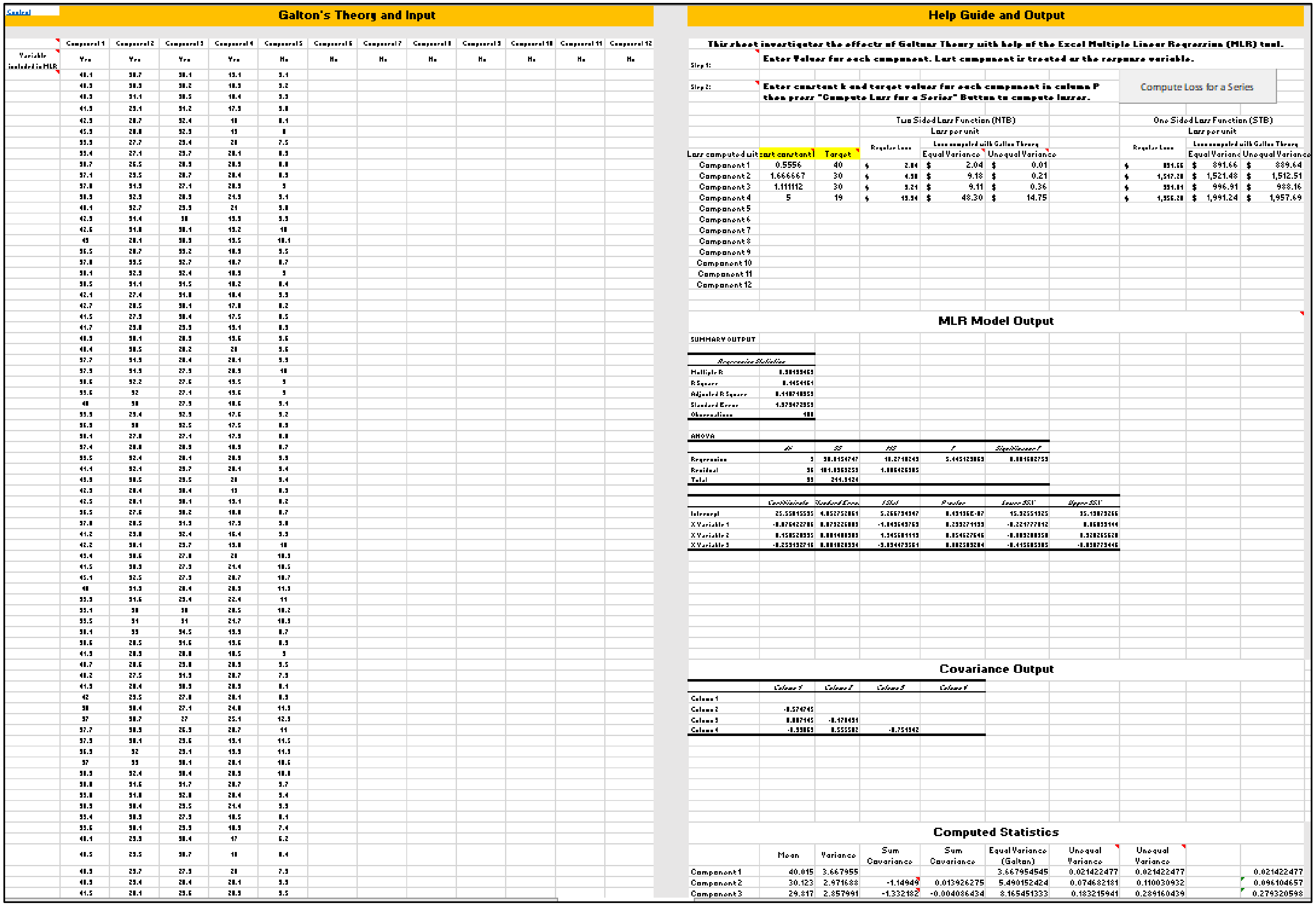


Figure 36. Excel simulation tool; spreadsheet 6 – Galton's Theory.

Content

Quantifiying the Loss of Quality Characteristic Variation for a Series System
Results

Moisture Content (Woody Residue 2018-07-23

Total Loss:

The variation in this series system causes a financial loss after shipment of the products in amount of:

Total Loss	Total Loss after Galton
2335.22 $	**8398.99 $**

Comment: This amount isolates the occuring variation on each individual component.

This amount considers the influence of the variation of each component on each other in a series system.

Individual Loss per Component:

Component	Loss Function Type	Loss per unit in [$/%²]	Galton Loss per unit in [$/%²]	Total Loss in [$]	Total Loss after Galton in [$]
Harvest / Collectior	NTB	2.02	2.02	201.91	201.91
Transport	NTB	4.93	10.99	493.19	1099.49
Drying	NTB	3.19	10.52	318.67	1051.91
Densification	NTB	13.21	60.46	1321.45	6045.68

Further Comments:

Figure 37. Excel simulation tool; spreadsheet 7 – summary.

The following code was developed for the bootstrap simulation from spreadsheet 3 and the graphical output from spreadsheet 4.

```
Private Sub Do_Simulation_Click()
' Clears Content of Sheets for next Simulation Execution
Worksheets("Computations").Range("B3:M100").Clear
Worksheets("Simulation_Output").Range("A4:DZ16").ClearContents
Worksheets("Summary").Range("A21:H37").ClearContents
On Error Resume Next
Worksheets("Simulation_Output").ChartObjects.Delete
On Error GoTo 0
' Delete cell color
Worksheets("Simulation_Output").Cells.Interior.ColorIndex = 0
' Declaration of the variables for the simulation
' Permanent user input variables
Dim num_com As Integer                              ' Number of components selected
Dim boot_iter As Long, boot_iter_str As String      ' Number of bootstrap iterations selected
Dim char_unit As String                             ' Unit of characteristic
Dim currency_unit As String                         ' Currency selected
Dim char_name As String                             ' Name of characteristic
' Loop variables
Dim m_col As Integer                                ' mth column of original sample
Dim m_asy_col As Integer                            ' mth column of original sample for asymmetric cases
Dim n_col As Integer                                ' nth column of original sample (needed to compute covariance)
Dim k As Long, k_str As String                      ' General loop variable
Dim boot_j As Long                                  ' Bootstrap iteration variable
Dim resa_j As Integer                               ' Resampling iterations based on bootstrap sample size
Dim i As Integer                                    ' Loop variable to determine the sample size for upper and lower dist
for NTB
Dim dist As Integer                                 ' Distance between components displayed on sheet
                                                    "Simulation_Output"
'Single user input for components of column m
Dim com_name As String                              ' Name of component
Dim loss_type As String                             ' Loss Function type
Dim target As Double                                ' Target
Dim USL As Double, LSL As Double                    ' USL/LSL of Component
Dim Loss_USL As Double, Loss_LSL As Double          ' Loss at spec limit
' Sample size / range Originial / Bootstrap
Dim size_boot As Integer                            ' Bootstrap sample size (Minimum of all original sample data sets)
Dim size_boot_up As Integer                         ' Bootstrap sample size for greater than target
Dim size_boot_low As Integer                        ' Bootstrap sample size for lower than target
Dim size_orig_input() As Variant                    ' Array which holds original sample sizes to determine minimum
Dim size_orig_m As Integer, rng_orig_m As Range     ' Size and Range for column m
Dim size_orig_n As Integer, rng_orig_n As Range     ' Size and Range for column n
' Computation Variables
Dim constant_k As Double                            ' constant k for NTB / STB / LTB
Dim cons_k_up As Double, cons_k_low As Double       ' Check for Asymmetry a. necessary for NTB asymmetric cases
Dim count_asym As Integer                           ' Counter for Asymmetry
Dim mean() As Double, Variance() As Double          ' Arrays which hold all means / variances from each bootstrap sample
Dim g_Mean As Double, g_Variance As Double          ' Grand Mean and Variance for column m
Dim g_Covariance As Double                          ' Grand covariance between mth and nth column
Dim sum_var As Double, sum_covar As Double          ' Sum of variance and covariance for mth component
Dim var_gal As Double                               ' Variance by galton's theory (series system)
Dim sum_larger As Double, g_Loss_larger As Double
Dim var_orig As Double
Dim Loss_orig As Double, Loss_gal As Double         ' Computed loss for components with and without Galton
Dim Loss_larger() As Double
' Variables for asymmetric cases
Dim count_upper As Integer                          ' Counts values in original sample above the target
Dim count_lower As Integer                          ' Counts values in original sample below the target
Dim Mean_upper() As Double                          ' Array which holds bootsrap statistic mean for USL side
Dim Mean_lower() As Double                          ' Array which holds bootsrap statistic mean for LSL side
Dim Variance_upper() As Double                      ' Array which holds bootsrap statistic variance for USL side
Dim Variance_lower() As Double                      ' Array which holds bootsrap statistic variance for LSL side
Dim m_count As Integer, n_count As Integer          ' counts for array above / below target
Dim g_mean_up As Double, g_mean_low As Double       ' Grand mean for upper and lower distribution
Dim g_var_up As Double, g_var_low As Double         ' Grand variance for upper and lower distribution
Dim g_covar_up As Double, g_covar_low As Double     ' Grand covariances for upper and lower distribution
Dim sum_var_up As Double, sum_var_low As Double     ' Sum of variances for upper and lower side
Dim sum_covar_up As Double, sum_covar_low As Double ' sum of covariances for upper and lower side
```

```
Dim var_gal_up As Double, var_gal_low As Double        ' Galton variance for upper and lower side
Dim var_orig_up As Double, var_orig_low As Double      ' Normal variance for upper and lower side
Dim Loss_orig_up As Double, Loss_orig_low As Double    ' Normal loss for NTB asymmetric cases
Dim Loss_gal_up As Double, Loss_gal_low As Double      ' Galton loss for NTB asymmetric cases
' Variables necessary to make graphs
Dim max_val_m As Double, min_val_m As Double           ' Min/Max value of column m
Dim bin_width As Double, places_array As Integer       ' Bin witdh for histograms, Number of places in array
Dim x_values() As Double, Loss_values() As Double      ' Arrays for x(characteristic) and y(loss) values
Dim step_value As Double, num_bins                     ' Each value in the arrays, determined number of bins
Dim places_array_min As Integer
Dim places_array_max As Integer
' Array for frequency per bin; Array for central value to get curve of original sample / bootstrap distribution
Dim freq_array() As Double, ctr_his_arr() As Double    ' Arrays which contain values to create histogram / sample distribution
Dim freq_histo_orig() As Variant                       ' Array which hold values from original sample
Dim freq_histo_boot() As Variant                       ' Array which holds all values from all bootstrap samples
Dim Orig_Loss As Chart, Boot_Loss As Chart             ' Defined charts
Dim boot_val_arr() As Variant, boot_val_arr_LTB As Variant     ' LTB arrays
Dim boot_val_count As Long, boot_val_count_LTB As Long         ' Count necessary to fill arrays
Dim boot_dist_arr_size As Long                         ' Array size determined through bootstrap iterations * bootstrap sample size
Dim x_values_max() As Double                           ' Array which holds x-values to calculate the loss function
Dim Loss_values_max() As Double                        ' Array which holds computed loss for each x-values
Dim calc_target As Double                              ' Needed to compute the loss function for asymmetric case for upper side
Dim Total_loss_series As Double                        ' Total loss of all components
Dim Total_Gal_loss_series As Double                    ' Total loss of all components after galton
Dim sum_row As Integer                                 ' Variable which counts the rows
' Get permanent User Input
num_com = Worksheets("User_Input").Range("F21").Value
boot_iter = Worksheets("User_Input").Range("F30").Value
boot_iter_str = boot_iter
char_unit = Worksheets("User_Input").Range("F17").Value
currency_unit = Worksheets("User_Input").Range("F18").Value
char_name = Worksheets("User_Input").Range("F16").Value
' Setting up Summary Sheet
With Worksheets("Summary")
   .Range("C11").Value = currency_unit
   .Range("G11").Value = currency_unit
End With
' Setting up "Computations" Sheet
For k = 1 To num_com
   k_str = k
   Worksheets("Computations").Cells(3, 2 + k).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(9 + k, 2).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(3, 2 + k).EntireColumn.AutoFit
Next k
With Worksheets("Computations")
   .Range("B4").Value = "Symmetry Check"
   .Range("B5").Value = "Grand Mean"
   .Range("B6").Value = "Grand Variance"
   .Range("B7").Value = "Sum Covariance"
   .Range("B8").Value = "Galton Variance"
   .Range("B9").Value = "Covariance Matrix"
End With
' Determining bootstrap sample size
' Due to unknown original sample sizes for each component the simulation
' determines the smallest original sample size and sets it as the bootstrap
' sample size
ReDim size_orig_input(1 To num_com)
For m_col = 1 To num_com
   size_orig_m = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + m_col), Cells(5000, 9 + m_col)))
   size_orig_input(m_col) = size_orig_m
   'Prints size original samples to sensitivity sheet --> needed for possible computations
   Worksheets("Sensitivity_Analysis").Cells(3, 21 + m_col).Value = size_orig_m
Next m_col
size_boot = WorksheetFunction.min(size_orig_input)
```

```
'Check for asymmetry
count_asym = 0
sum_row = 0
For k = 1 To num_com
    If Worksheets("User_Input").Cells(7, 9 + k).Value = "Nominal-the-best" Then
        target = Worksheets("User_Input").Cells(8, 9 + k).Value
        USL = Worksheets("User_Input").Cells(9, 9 + k).Value
        LSL = Worksheets("User_Input").Cells(10, 9 + k).Value
        Loss_USL = Worksheets("User_Input").Cells(11, 9 + k).Value
        Loss_LSL = Worksheets("User_Input").Cells(12, 9 + k).Value
        cons_k_up = (Loss_USL / ((USL - target) ^ 2))
        cons_k_low = (Loss_LSL / ((target - LSL) ^ 2))
        If cons_k_up = cons_k_low Then
            Worksheets("Computations").Cells(4, 2 + k).Value = "Yes"
            count_asym = count_asym + 0
        Else
            Worksheets("Computations").Cells(4, 2 + k).Value = "No"
            count_asym = count_asym + 1
        End If
    Else
        Worksheets("Computations").Cells(4, 2 + k).Value = "Yes"
        count_asym = count_asym + 0
    End If
Next k
' First For-Loop to get data from first column m
For m_col = 1 To num_com
' Check for Loss Function type
    Select Case Worksheets("User_Input").Cells(7, 9 + m_col)
    ' Get individual component user input
        Case "Nominal-the-best"
            com_name = Worksheets("User_Input").Cells(6, 9 + m_col).Value
            loss_type = Worksheets("User_Input").Cells(7, 9 + m_col).Value
            target = Worksheets("User_Input").Cells(8, 9 + m_col).Value
            USL = Worksheets("User_Input").Cells(9, 9 + m_col).Value
            LSL = Worksheets("User_Input").Cells(10, 9 + m_col).Value
            Loss_USL = Worksheets("User_Input").Cells(11, 9 + m_col).Value
            Loss_LSL = Worksheets("User_Input").Cells(12, 9 + m_col).Value
            ' Computing constant k for upper and lower side of NTB
            ' Check for Symmetry with Spec / Loss --> constant k
            cons_k_up = (Loss_USL / ((USL - target) ^ 2))
            cons_k_low = (Loss_LSL / ((target - LSL) ^ 2))
            If cons_k_up <> cons_k_low Then
                count_asym = 1
            Else
                constant_k = cons_k_up
                count_asym = 0
            End If
        Case "Smaller-the-better"
            com_name = Worksheets("User_Input").Cells(6, 9 + m_col).Value
            loss_type = Worksheets("User_Input").Cells(7, 9 + m_col).Value
            target = Worksheets("User_Input").Cells(8, 9 + m_col).Value
            'Worksheets("User_Input").Cells(9, 9 + m_col).Value = "Does not apply"
            'Worksheets("User_Input").Cells(10, 9 + m_col).Value = "Does not apply"
            Loss_USL = Worksheets("User_Input").Cells(11, 9 + m_col).Value
            'Worksheets("User_Input").Cells(12, 9 + m_col).Value = "Does not apply"
            'Computing constant k
            constant_k = (Loss_USL / (target ^ 2))
            count_asym = 0
        Case "Larger-the-better"
            com_name = Worksheets("User_Input").Cells(6, 9 + m_col).Value
            loss_type = Worksheets("User_Input").Cells(7, 9 + m_col).Value
            target = Worksheets("User_Input").Cells(8, 9 + m_col).Value
            'Worksheets("User_Input").Cells(9, 9 + m_col).Value = "Does not apply"
            'Worksheets("User_Input").Cells(10, 9 + m_col).Value = "Does not apply"
            Loss_USL = Worksheets("User_Input").Cells(11, 9 + m_col).Value
            'Worksheets("User_Input").Cells(12, 9 + m_col).Value = "Does not apply"
            constant_k = (Loss_USL * (target ^ 2))
```

```
        count_asym = 0
    End Select

    ' Checks if quality loss function is asymmetric for NTB (checks for all cases due to syntax)
    ' Setting size and range of original samples from column m and n depending on the size
    size_orig_m = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + m_col), Cells(5000, 9 + m_col)))
    Set rng_orig_m = Worksheets("User_Input").Range(Cells(13, 9 + m_col), Cells(size_orig_m + 12, 9 + m_col))
    ' First Bootstrap simulation for column k
    'Starting bootstrap simulation by reallocating memory of mean/variance/
    ReDim mean(1 To boot_iter)
    ReDim Variance(1 To boot_iter)
    ReDim Loss_larger(1 To boot_iter)
    ' Computing array size of bootstrap-distribution and reallocate bootstrap distribution array
    boot_dist_arr_size = boot_iter * size_orig_m
    ReDim boot_val_arr(1 To boot_dist_arr_size)
    ReDim boot_val_arr_LTB(1 To boot_dist_arr_size)
    boot_val_count = 1
    boot_val_count_LTB = 1
    'First Bootstrap loop for column m
    For boot_j = 1 To boot_iter
      ReDim val_res_m(1 To size_boot)
      ' Second For-loop of bootstrapping --> randomly drawing values with resampling
      For resa_j = 1 To size_boot
        val_res_m(resa_j) = WorksheetFunction.Index(rng_orig_m, size_orig_m * Rnd() + 1)
      Next resa_j
      ' For-loop to add all drawn values to one array --> bootstrap distribution
      For i = 1 To size_orig_m
        boot_val_arr(boot_val_count) = val_res_m(i)
        boot_val_count = boot_val_count + 1
      Next i
      ' Computing mean and variance for each bootstrap sample
      mean(boot_j) = WorksheetFunction.Average(val_res_m)
      Variance(boot_j) = WorksheetFunction.Var_S(val_res_m)
      ' Larger-the-better computation
      If Worksheets("User_Input").Cells(7, 9 + m_col).Value = "Larger-the-better" Then
        ReDim val_res_m_LTB(1 To size_orig_m)
        For resa_j = 1 To size_orig_m
          val_res_m_LTB(resa_j) = WorksheetFunction.Index(rng_orig_m, size_orig_m * Rnd() + 1)
        Next resa_j
        sum_larger = 0
        For i = 1 To size_orig_m
          sum_larger = sum_larger + (1 / (val_res_m_LTB(i) ^ 2))
        Next i
        ' Computing loss for one bootstrap sample
        Loss_larger(boot_j) = constant_k * (1 / size_orig_m) * sum_larger
        For i = 1 To size_orig_m
          boot_val_arr_LTB(boot_val_count_LTB) = val_res_m_LTB(i)
          boot_val_count_LTB = boot_val_count_LTB + 1
        Next i
      End If
    Next boot_j
    Call Histogram_means(Loss_larger, dist, boot_iter_str)
    'Computing the loss for Larger-the-better cases
    g_Loss_larger = 0
    If Worksheets("User_Input").Cells(7, 9 + m_col).Value = "Larger-the-better" Then
      g_Loss_larger = WorksheetFunction.Average(Loss_larger)
    End If
    ' Computing/printing grand statistics for mean, variance, and covariance
    g_Mean = WorksheetFunction.Average(mean)
    Worksheets("Computations").Cells(5, 2 + m_col).Value = g_Mean
    g_Variance = WorksheetFunction.Average(Variance)
    Worksheets("Computations").Cells(6, 2 + m_col).Value = g_Variance
    ' Second For-loop, necessary for covariance computations
    For n_col = 2 To num_com
      size_orig_n = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + n_col), Cells(5000, 9 + n_col)))
      Set rng_orig_n = Worksheets("User_Input").Range(Cells(13, 9 + n_col), Cells(size_orig_n + 12, 9 + n_col))
      ' Starting bootstrap simulation by reallocating memory of covariance
```

```
For n_col = 2 To num_com
   size_orig_n = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + n_col), Cells(5000, 9 + n_col)))
   Set rng_orig_n = Worksheets("User_Input").Range(Cells(13, 9 + n_col), Cells(size_orig_n + 12, 9 + n_col))
   ' Starting bootstrap simulation by reallocating memory of covariance
   ReDim Covariance(1 To boot_iter)
   ' Nested For-loops; First For-loop loops through the number of bootstrap iterations
   For boot_j = 1 To boot_iter
      ReDim val_res_n(1 To size_boot)
      ' Second For-loop loops through the number of resample draws
      For resa_j = 1 To size_boot
         ' Randomly takes values from original sample to bootstrap with resampling
         ' Procedure assumes non-parametric distribution of original sample
         val_res_n(resa_j) = WorksheetFunction.Index(rng_orig_n, size_orig_n * Rnd() + 1)
      Next resa_j
      ' Computing covariance for each bootstrap sample
      ' If then statement for computing covariances between each data set / column
      If m_col < num_com And m_col < n_col Then
         Covariance(boot_j) = WorksheetFunction.Covariance_S(val_res_m, val_res_n)
         Worksheets("Computations").Cells(boot_j + 1, 10 + m_col).Value = Covariance(boot_j)
      End If

   Next boot_j
   ' If then statement for computing covariances between each data set / column
   If m_col < num_com And m_col < n_col Then
      g_Covariance = WorksheetFunction.Average(Covariance)
      Worksheets("Computations").Cells(m_col + 9, n_col + 2).Value = g_Covariance
   End If
Next n_col
' Getting Values and computing statistics for asymmetric case if applicable
' Checks if asymmetric computations are necessary
If count_asym >= 1 Then
' Setting up "Computations" Sheet for asymmetric cases
For k = 1 To num_com
   k_str = k
   Worksheets("Computations").Cells(30, 2 + k).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(35 + k, 2).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(30, 2 + k).EntireColumn.AutoFit
   Worksheets("Computations").Cells(50, 2 + k).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(55 + k, 2).Value = "Component" + " " + k_str
   Worksheets("Computations").Cells(50, 2 + k).EntireColumn.AutoFit
Next k
With Worksheets("Computations")
   .Range("B30").Value = "Upper Side of NTB"
   .Range("B31").Value = "Grand Mean"
   .Range("B32").Value = "Grand Variance"
   .Range("B33").Value = "Sum Covariance"
   .Range("B34").Value = "Galton Variance"
   .Range("B35").Value = "Covariance Matrix"
   .Range("B50").Value = "Lower Side of NTB"
   .Range("B51").Value = "Grand Mean"
   .Range("B52").Value = "Grand Variance"
   .Range("B53").Value = "Sum Covariance"
   .Range("B54").Value = "Galton Variance"
   .Range("B55").Value = "Covariance Matrix"
End With
'Reallocation of memory for both arrays
ReDim upper_size(1 To num_com)
ReDim lower_size(1 To num_com)
' Finding the smallest original sample size for upper and lower side of curve
' Necessary for bootstrap simulation - bootsrap sample size is the lowest of all
For k = 1 To num_com
   count_upper = 0
   count_lower = 0
   size_orig_m = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + k), Cells(5000, 9 + k)))
   For i = 1 To size_orig_m
      If Cells(12 + i, 9 + k) > Cells(8, 9 + k) Then
         count_upper = count_upper + 1
```

```
        ReDim Covariance(1 To boot_iter)
        ' Nested For-loops; First For-loop loops through the number of bootstrap iterations
        For boot_j = 1 To boot_iter
            ReDim val_res_n(1 To size_boot)
            ' Second For-loop loops through the number of resample draws
            For resa_j = 1 To size_boot
                ' Randomly takes values from original sample to bootstrap with resampling
                ' Procedure assumes non-parametric distribution of original sample
                val_res_n(resa_j) = WorksheetFunction.Index(rng_orig_n, size_orig_n * Rnd() + 1)
            Next resa_j
            ' Computing covariance for each bootstrap sample
            ' If then statement for computing covariances between each data set / column
            If m_col < num_com And m_col < n_col Then
                Covariance(boot_j) = WorksheetFunction.Covariance_S(val_res_m, val_res_n)
            End If

        Next boot_j
        ' If then statement for computing covariances between each data set / column
        If m_col < num_com And m_col < n_col Then
            g_Covariance = WorksheetFunction.Average(Covariance)
            Worksheets("Computations").Cells(m_col + 9, n_col + 2).Value = g_Covariance
        End If
    Next n_col
    ' Getting Values and computing statistics for asymmetric case if applicable
    ' Checks if asymmetric computations are necessary
    If count_asym >= 1 Then
    ' Setting up "Computations" Sheet for asymmetric cases
    For k = 1 To num_com
        k_str = k
        Worksheets("Computations").Cells(30, 2 + k).Value = "Component" + " " + k_str
        Worksheets("Computations").Cells(35 + k, 2).Value = "Component" + " " + k_str
        Worksheets("Computations").Cells(30, 2 + k).EntireColumn.AutoFit
        Worksheets("Computations").Cells(50, 2 + k).Value = "Component" + " " + k_str
        Worksheets("Computations").Cells(55 + k, 2).Value = "Component" + " " + k_str
        Worksheets("Computations").Cells(50, 2 + k).EntireColumn.AutoFit
    Next k
    With Worksheets("Computations")
        .Range("B30").Value = "Upper Side of NTB"
        .Range("B31").Value = "Grand Mean"
        .Range("B32").Value = "Grand Variance"
        .Range("B33").Value = "Sum Covariance"
        .Range("B34").Value = "Galton Variance"
        .Range("B35").Value = "Covariance Matrix"
        .Range("B50").Value = "Lower Side of NTB"
        .Range("B51").Value = "Grand Mean"
        .Range("B52").Value = "Grand Variance"
        .Range("B53").Value = "Sum Covariance"
        .Range("B54").Value = "Galton Variance"
        .Range("B55").Value = "Covariance Matrix"
    End With
    'Reallocation of memory for both arrays
    ReDim upper_size(1 To num_com)
    ReDim lower_size(1 To num_com)
    ' Finding the smallest original sample size for upper and lower side of curve
    ' Necessary for bootstrap simulation - bootsrap sample size is the lowest of all
    For k = 1 To num_com
        count_upper = 0
        count_lower = 0
        size_orig_m = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + k), Cells(5000, 9 + k)))
        For i = 1 To size_orig_m
            If Cells(12 + i, 9 + k) > Cells(8, 9 + k) Then
                count_upper = count_upper + 1
            ElseIf Cells(12 + i, 9 + k) < Cells(8, 9 + k) Then
                count_lower = count_lower + 1
            End If
        Next i
        upper_size(k) = count_upper
```

```
    lower_size(k) = count_lower
  Next k
  size_boot_up = WorksheetFunction.min(upper_size)
  size_boot_low = WorksheetFunction.min(lower_size)
  ' Nested For-loop to get all values for upper and lower side of target
  sum_covar_up = 0
  sum_covar_low = 0
  var_gal_up = 0
  var_gal_low = 0
  sum_var_up = 0
  sum_var_low = 0
  For m_asy_col = 1 To num_com
    ' Checks original sample size for all data
    size_orig_m = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + m_asy_col), Cells(5000, 9 +
m_asy_col)))
    ' Reallocation of memory for the array for column m
    ReDim orig_up_m(1 To upper_size(m_asy_col))
    ReDim orig_low_m(1 To lower_size(m_asy_col))
    ' Initializing of upper and lower array of one column of values to 1
    m_count = 1
    n_count = 1
    ' Checks values whether bigger or smaller than target and ads values to certain array
    For i = 1 To size_orig_m
      If Cells(12 + i, 9 + m_asy_col) > Cells(8, 9 + m_asy_col) Then
        orig_up_m(m_count) = Cells(12 + i, 9 + m_asy_col).Value
        m_count = m_count + 1
      ElseIf Cells(12 + i, 9 + m_asy_col) < Cells(8, 9 + m_asy_col) Then
        orig_low_m(n_count) = Cells(12 + i, 9 + m_asy_col).Value
        n_count = n_count + 1
      End If
    Next i
    ' Second for loop to get values for column n
    ' Reallocation of mean and varaince arrays
    ReDim Mean_upper(1 To boot_iter)
    ReDim Mean_lower(1 To boot_iter)
    ReDim Variance_upper(1 To boot_iter)
    ReDim Variance_lower(1 To boot_iter)
    For boot_j = 1 To boot_iter
      ' Reallocation bootstrap sample values for column m for upper / lower side
      ReDim val_res_up_m(1 To size_boot_up)
      ReDim val_res_low_m(1 To size_boot_low)
      ' Second For-loop to resample values from original sample to bootstrap sample
      ' Randomly takes values from original sample to bootstrap with resampling
      ' Procedure assumes non-parametric distribution of original sample
      For resa_j = 1 To size_boot_up
        val_res_up_m(resa_j) = WorksheetFunction.Index(orig_up_m, upper_size(m_asy_col) * Rnd() + 1)
      Next resa_j
      For resa_j = 1 To size_boot_low
        val_res_low_m(resa_j) = WorksheetFunction.Index(orig_low_m, lower_size(m_asy_col) * Rnd() + 1)
      Next resa_j
      ' Computing statistic for each individual bootstrap sample
      Mean_upper(boot_j) = WorksheetFunction.Average(val_res_up_m)
      Mean_lower(boot_j) = WorksheetFunction.Average(val_res_low_m)
      Variance_upper(boot_j) = WorksheetFunction.Var_S(val_res_up_m)
      Variance_lower(boot_j) = WorksheetFunction.Var_S(val_res_low_m)
    Next boot_j
    ' Computing/printing grand statistics for mean, variance, and covariance
    g_mean_up = WorksheetFunction.Average(Mean_upper)
    g_mean_low = WorksheetFunction.Average(Mean_lower)
    g_var_up = WorksheetFunction.Average(Variance_upper)
    g_var_low = WorksheetFunction.Average(Variance_lower)
    Worksheets("Computations").Cells(31, 2 + m_asy_col).Value = g_mean_up
    Worksheets("Computations").Cells(51, 2 + m_asy_col).Value = g_mean_low
    Worksheets("Computations").Cells(32, 2 + m_asy_col).Value = g_var_up
    Worksheets("Computations").Cells(52, 2 + m_asy_col).Value = g_var_low
    For n_col = 2 To num_com
      'Checks original sample size for whole data set
```

```
        size_orig_n = WorksheetFunction.CountA(Worksheets("User_Input").Range(Cells(13, 9 + n_col), Cells(5000, 9 + n_col)))
        ' Reallocation of memory for the array for column n
        ReDim orig_up_n(1 To upper_size(n_col))
        ReDim orig_low_n(1 To lower_size(n_col))
        ' Initializing upper and lower array starting position
        m_count = 1
        n_count = 1
        ' Checks values whether bigger or smaller than target and adds values to certain array
        For i = 1 To size_orig_n
            If Cells(12 + i, 9 + n_col) > Cells(8, 9 + n_col) Then
                orig_up_n(m_count) = Cells(12 + i, 9 + n_col).Value
                m_count = m_count + 1
            ElseIf Cells(12 + i, 9 + n_col) < Cells(8, 9 + n_col) Then
                orig_low_n(n_count) = Cells(12 + i, 9 + n_col).Value
                n_count = n_count + 1
            End If
        Next i
        ' Reallocation of covariance arrays
        ReDim Covariance_upper(1 To boot_iter)
        ReDim Covariance_lower(1 To boot_iter)
        ' Nested for-loop for bootstrap simulation
        ' First For-loop iteration of bootstrap simulation
        For boot_j = 1 To boot_iter
            ' Reallocation bootstrap sample values for column n for upper / lower side
            ReDim val_res_up_n(1 To size_boot_up)
            ReDim val_res_low_n(1 To size_boot_low)
            ' Second For-loop to resample values from original sample to bootstrap sample
            ' Randomly takes values from original sample to bootstrap with resampling
            ' Procedure assumes non-parametric distribution of original sample
            For resa_j = 1 To size_boot_up
                val_res_up_n(resa_j) = WorksheetFunction.Index(orig_up_n, upper_size(n_col) * Rnd() + 1)
            Next resa_j
            For resa_j = 1 To size_boot_low
                val_res_low_n(resa_j) = WorksheetFunction.Index(orig_low_n, lower_size(n_col) * Rnd() + 1)
            Next resa_j
            If m_asy_col < num_com And m_asy_col < n_col Then
                Covariance_upper(boot_j) = WorksheetFunction.Covariance_S(val_res_up_m, val_res_up_n)
                Covariance_lower(boot_j) = WorksheetFunction.Covariance_S(val_res_low_m, val_res_low_n)
            End If
        Next boot_j
        ' If-Then statement prevents computation of covariance with columns with itself and repetitions
        If m_asy_col < num_com And m_asy_col < n_col Then
            g_covar_up = WorksheetFunction.Average(Covariance_upper)
            Worksheets("Computations").Cells(m_asy_col + 35, n_col + 2).Value = g_covar_up
            g_covar_low = WorksheetFunction.Average(Covariance_lower)
            Worksheets("Computations").Cells(m_asy_col + 55, n_col + 2).Value = g_covar_low
        End If
    Next n_col
    ' Computing normal variance, total sum of all covariances, and variance for galton
    sum_var_up = sum_var_up + Worksheets("Computations").Cells(32, 2 + m_asy_col).Value
    sum_var_low = sum_var_low + Worksheets("Computations").Cells(52, 2 + m_asy_col).Value
    ' Computing the correct sums of covariance
    For i = 2 To num_com
        sum_covar_up = sum_covar_up + Worksheets("Computations").Cells(34 + i, 3 + m_asy_col).Value
        sum_covar_low = sum_covar_low + Worksheets("Computations").Cells(54 + i, 3 + m_asy_col).Value
    Next i
    ' Printing covariance in certain cells for later computation of loss
    If m_asy_col < num_com Then
        Worksheets("Computations").Cells(33, 3 + m_asy_col).Value = sum_covar_up
        Worksheets("Computations").Cells(53, 3 + m_asy_col).Value = sum_covar_low
    End If
    ' Computing the total variance after Galton's Theory for a series system
    var_gal_up = sum_var_up + Worksheets("Computations").Cells(33, 2 + m_asy_col).Value * 2
    var_gal_low = sum_var_low + Worksheets("Computations").Cells(53, 2 + m_asy_col).Value * 2
    Worksheets("Computations").Cells(34, 2 + m_asy_col).Value = var_gal_up
    Worksheets("Computations").Cells(54, 2 + m_asy_col).Value = var_gal_low
Next m_asy_col
```

```
End If
' Computation of the variance for each step based for a series system (Galton)
' For loop to move through the columns on Sheet "Computations"
sum_var = sum_var + Worksheets("Computations").Cells(6, 2 + m_col).Value
' For loop to sum covaraince
For i = 2 To num_com
   sum_covar = sum_covar + Worksheets("Computations").Cells(8 + i, 3 + m_col).Value
Next i
' Printing sum of covariances in correct column
If m_col < num_com Then
   Worksheets("Computations").Cells(7, 3 + m_col).Value = sum_covar
End If
' Computing the total variance after Galton's Theory for a series system
var_gal = sum_var + 2 * Worksheets("Computations").Cells(7, m_col + 2).Value
Worksheets("Computations").Cells(8, m_col + 2).Value = var_gal
' Computation of the loss for NTB and STB
If Worksheets("Computations").Cells(4, 2 + m_col).Value = "Yes" Then
   ' Getting grand mean, variance (normal), and galton variance from sheet "Computation" for loss calculation
   g_Mean = Worksheets("Computations").Cells(5, 2 + m_col).Value
   var_orig = Worksheets("Computations").Cells(6, 2 + m_col).Value
   var_gal = Worksheets("Computations").Cells(8, 2 + m_col).Value
   If Worksheets("User_Input").Cells(7, 9 + m_col).Value = "Nominal-the-best" Then
      ' Computing loss without series influence
      Loss_orig = constant_k * (var_orig + (g_Mean - target) ^ 2)
      ' Computing loss with series influence --> Galton
      Loss_gal = constant_k * (var_gal + (g_Mean - target) ^ 2)
   ' Calculation for Smaller-the-better Loss Function
   Else
      ' Computing loss without series influence
      Loss_orig = constant_k * (var_orig + g_Mean ^ 2)
      ' Computing loss with series influence
      Loss_gal = constant_k * (var_gal + g_Mean ^ 2)
   End If
Else
   g_mean_up = Worksheets("Computations").Cells(31, 2 + m_col).Value
   g_mean_low = Worksheets("Computations").Cells(51, 2 + m_col).Value
   var_orig_up = Worksheets("Computations").Cells(32, 2 + m_col).Value
   var_orig_low = Worksheets("Computations").Cells(52, 2 + m_col).Value
   var_gal_up = Worksheets("Computations").Cells(34, 2 + m_col).Value
   var_gal_low = Worksheets("Computations").Cells(54, 2 + m_col).Value
   ' Computing loss without series influence for asymmetric case
   Loss_orig_up = cons_k_up * (var_orig_up + (g_mean_up - target) ^ 2)
   Loss_orig_low = cons_k_low * (var_orig_low + (g_mean_low - target) ^ 2)
   ' Computing loss with series influence for symmetric case
   Loss_gal_up = cons_k_up * (var_gal_up + (g_mean_up - target) ^ 2)
   Loss_gal_low = cons_k_low * (var_gal_low + (g_mean_low - target) ^ 2)
End If
' Printing all results on Sheet "Simulation Output"
k_str = m_col
' General data for all loss functions
Dim r As Range
'Set r = Worksheets("Simulation_Output").Range(Cells(4, 3 + dist), Cells(5, 8 + dist))
'Set r = Worksheets("Simulation_Output").Range(Cells(1, 1), Cells(2, 2))
'r.Interior.Color = RGB(255, 0, 0)
With Worksheets("Simulation_Output")
   .Cells(4, 5 + dist).Value = com_name
   '.Range(Cells(4, 5 + dist), Cells(5, 8 + dist)).Merge
   .Cells(4, 3 + dist).Value = "Component" + " " + k_str
   '.Range(Cells(4, 3 + dist), Cells(4, 4 + dist)).Merge
   .Cells(5, 3 + dist).Value = loss_type
   '.Range(Cells(5, 3 + dist), Cells(5, 4 + dist)).Merge
   .Cells(7, 3 + dist).Value = "Input"
   .Cells(8, 3 + dist).Value = "Target"
   .Cells(8, 4 + dist).Value = target
   .Cells(8, 5 + dist).Value = char_unit
   .Cells(10, 3 + dist).Value = "Output"
   '.Cells(4, 5 + dist), Cells(5, 8 + dist).Interior.Color = RGB(255, 0, 0)
```

```
End With
' Setting up main parts of summary sheet
With Worksheets("Summary")
    .Range("A20").Value = "Component"
    .Range("B20").Value = "Loss Function Type"
    .Range("C20").Value = "Loss per unit in " + "[" + currency_unit + "/" + char_unit + "²" + "]"
    .Range("E20").Value = "Galton Loss per unit in " + "[" + currency_unit + "/" + char_unit + "²" + "]"
    .Range("F20").Value = "Total Loss in " + "[" + currency_unit + "]"
    .Range("H20").Value = "Total Loss after Galton in " + "[" + currency_unit + "]"
End With
' Setting up "Simulation_Output" sheet
Select Case Worksheets("User_Input").Cells(7, 9 + m_col).Value
    Case "Nominal-the-best"
        With Worksheets("Simulation_Output")
         .Cells(8, 6 + dist).Value = "USL"
         .Cells(8, 7 + dist).Value = USL
         .Cells(8, 8 + dist).Value = char_unit
         .Cells(9, 6 + dist).Value = "LSL"
         .Cells(9, 7 + dist).Value = LSL
         .Cells(9, 8 + dist).Value = char_unit
        End With
        If Worksheets("Computations").Cells(4, 2 + m_col).Value = "Yes" Then
            With Worksheets("Simulation_Output")
                .Cells(12, 3 + dist).Value = "k"
                .Cells(12, 5 + dist).Value = constant_k
                .Cells(12, 6 + dist).Value = currency_unit + "/" + char_unit + "²"
                .Cells(13, 3 + dist).Value = "Loss"
                .Cells(13, 4 + dist).Value = Round(Loss_orig, 2)
                .Cells(13, 5 + dist).Value = currency_unit + " per unit"
                .Cells(13, 6 + dist).Value = "Galton Loss"
                .Cells(13, 7 + dist).Value = Round(Loss_gal, 2)
                .Cells(13, 8 + dist).Value = currency_unit + " per unit"
                .Cells(14, 3 + dist).Value = "Total Loss"
                .Cells(14, 4 + dist).Value = Round(Loss_orig * size_orig_m, 2)
                .Cells(14, 5 + dist).Value = currency_unit
                .Cells(14, 6 + dist).Value = "Total Galton Loss"
                .Cells(14, 7 + dist).Value = Round(Loss_gal * size_orig_m, 2)
                .Cells(14, 8 + dist).Value = currency_unit
            End With
            With Worksheets("Summary")
                .Cells(21 + sum_row, 1).Value = com_name
                .Cells(21 + sum_row, 2).Value = "NTB"
                .Cells(21 + sum_row, 3).Value = Round(Loss_orig, 2)
                .Cells(21 + sum_row, 5).Value = Round(Loss_gal, 2)
                .Cells(21 + sum_row, 6).Value = Round(Loss_orig * size_orig_m, 2)
                .Cells(21 + sum_row, 8).Value = Round(Loss_gal * size_orig_m, 2)
            End With
                sum_row = sum_row + 1
        ElseIf Worksheets("Computations").Cells(4, 2 + m_col).Value = "No" Then
            With Worksheets("Simulation_Output")
                .Cells(11, 3 + dist).Value = "Lower Spec Side"
                .Cells(11, 6 + dist).Value = "Upper Spec Side"
                .Cells(12, 3 + dist).Value = "k"
                .Cells(12, 4 + dist).Value = cons_k_low
                .Cells(12, 5 + dist).Value = currency_unit + "/" + char_unit + "²"
                .Cells(13, 3 + dist).Value = "Loss"
                .Cells(13, 4 + dist).Value = Round(Loss_orig_low, 2)
                .Cells(13, 5 + dist).Value = currency_unit + " per unit"
                .Cells(14, 3 + dist).Value = "Galton Loss"
                .Cells(14, 4 + dist).Value = Round(Loss_gal_low, 2)
                .Cells(14, 5 + dist).Value = currency_unit + " per unit"
                .Cells(15, 3 + dist).Value = "Total Loss"
                .Cells(15, 4 + dist).Value = Round(Loss_orig_low * lower_size(m_col), 2)
                .Cells(15, 5 + dist).Value = currency_unit
                .Cells(16, 3 + dist).Value = "Total Galton Loss"
                .Cells(16, 4 + dist).Value = Round(Loss_gal_low * lower_size(m_col), 2)
                .Cells(16, 5 + dist).Value = currency_unit
```

```
            .Cells(12, 6 + dist).Value = "k"
            .Cells(12, 7 + dist).Value = cons_k_up
            .Cells(12, 8 + dist).Value = currency_unit + "/" + char_unit + "²"
            .Cells(13, 6 + dist).Value = "Loss"
            .Cells(13, 7 + dist).Value = Round(Loss_orig_up, 2)
            .Cells(13, 8 + dist).Value = currency_unit + " per unit"
            .Cells(14, 6 + dist).Value = "Galton Loss"
            .Cells(14, 7 + dist).Value = Round(Loss_gal_up, 2)
            .Cells(14, 8 + dist).Value = currency_unit + " per unit"
            .Cells(15, 6 + dist).Value = "Total Loss"
            .Cells(15, 7 + dist).Value = Round(Loss_orig_up * upper_size(m_col), 2)
            .Cells(15, 8 + dist).Value = currency_unit
            .Cells(16, 6 + dist).Value = "Total Galton Loss"
            .Cells(16, 7 + dist).Value = Round(Loss_gal_up * upper_size(m_col), 2)
            .Cells(16, 8 + dist).Value = currency_unit
        End With
        With Worksheets("Summary")
            .Cells(21 + sum_row, 1).Value = com_name
            .Cells(21 + sum_row, 2).Value = "NTB Upper Side"
            .Cells(21 + sum_row, 3).Value = Round(Loss_orig_up, 2)
            .Cells(21 + sum_row, 5).Value = Round(Loss_gal_up, 2)
            .Cells(21 + sum_row, 6).Value = Round(Loss_orig_up * size_orig_m, 2)
            .Cells(21 + sum_row, 8).Value = Round(Loss_gal_up * size_orig_m, 2)
        End With
            sum_row = sum_row + 1
        With Worksheets("Summary")
            .Cells(21 + sum_row, 2).Value = "NTB Lower Side"
            .Cells(21 + sum_row, 3).Value = Round(Loss_orig_low, 2)
            .Cells(21 + sum_row, 5).Value = Round(Loss_gal_low, 2)
            .Cells(21 + sum_row, 6).Value = Round(Loss_orig_low * size_orig_m, 2)
            .Cells(21 + sum_row, 8).Value = Round(Loss_gal_low * size_orig_m, 2)
        End With
            sum_row = sum_row + 1
    End If
Case "Smaller-the-better"
    With Worksheets("Simulation_Output")
        .Cells(8, 6 + dist).Value = "Loss at Target"
        .Cells(8, 7 + dist).Value = Loss_USL
        .Cells(8, 8 + dist).Value = currency_unit
        .Cells(12, 3 + dist).Value = "k"
        .Cells(12, 4 + dist).Value = constant_k
        .Cells(12, 5 + dist).Value = currency_unit + "/" + char_unit + "²"
        .Cells(13, 3 + dist).Value = "Loss"
        .Cells(13, 4 + dist).Value = Round(Loss_orig, 2)
        .Cells(13, 5 + dist).Value = currency_unit + " per unit"
        .Cells(13, 6 + dist).Value = "Galton Loss"
        .Cells(13, 7 + dist).Value = Round(Loss_gal, 2)
        .Cells(13, 8 + dist).Value = currency_unit + " per unit"
        .Cells(14, 3 + dist).Value = "Total Loss"
        .Cells(14, 4 + dist).Value = Round(Loss_orig * size_orig_m, 2)
        .Cells(14, 5 + dist).Value = currency_unit
        .Cells(14, 6 + dist).Value = "Total Galton Loss"
        .Cells(14, 7 + dist).Value = Round(Loss_gal * size_orig_m, 2)
        .Cells(14, 8 + dist).Value = currency_unit
    End With
    With Worksheets("Summary")
        .Cells(21 + sum_row, 1).Value = com_name
        .Cells(21 + sum_row, 2).Value = "STB"
        .Cells(21 + sum_row, 3).Value = Round(Loss_orig, 2)
        .Cells(21 + sum_row, 5).Value = Round(Loss_gal, 2)
        .Cells(21 + sum_row, 6).Value = Round(Loss_orig * size_orig_m, 2)
        .Cells(21 + sum_row, 8).Value = Round(Loss_gal * size_orig_m, 2)
    End With
        sum_row = sum_row + 1
Case "Larger-the-better"
    With Worksheets("Simulation_Output")
        .Cells(8, 6 + dist).Value = "Loss at Target"
```

```
            .Cells(8, 7 + dist).Value = Loss_USL
            .Cells(8, 8 + dist).Value = currency_unit
            .Cells(12, 4 + dist).Value = "k"
            .Cells(12, 5 + dist).Value = constant_k
            .Cells(12, 6 + dist).Value = currency_unit + "/" + char_unit + "²"
            .Cells(13, 4 + dist).Value = "Average unit loss"
            .Cells(13, 6 + dist).Value = Round(g_Loss_larger, 2)
            .Cells(13, 7 + dist).Value = currency_unit + " per unit"
            .Cells(14, 4 + dist).Value = "Total Loss"
            .Cells(14, 6 + dist).Value = Round(g_Loss_larger * size_orig_m, 2)
            .Cells(14, 7 + dist).Value = currency_unit
        End With
        With Worksheets("Summary")
            .Cells(21 + sum_row, 1).Value = com_name
            .Cells(21 + sum_row, 2).Value = "LTB"
            .Cells(21 + sum_row, 3).Value = Round(g_Loss_larger, 2)
            .Cells(21 + sum_row, 6).Value = Round(g_Loss_larger * size_orig_m, 2)
        End With
            sum_row = sum_row + 1
End Select
' Setting up graphs and charts in Sheet "Simulation_Output"
' Determining min and max value of original sample size to get range and
' places of array which holds x-values (x-axis) and loss-values (y-values)
Select Case Worksheets("User_Input").Cells(7, 9 + m_col).Value
    Case "Nominal-the-best"
        If Worksheets("Computations").Cells(4, 2 + m_col).Value = "Yes" Then
            ' Determining max and min value of original sample
            max_val_m = WorksheetFunction.max(rng_orig_m)
            min_val_m = WorksheetFunction.min(rng_orig_m)
            ' Computing bin width for histogram / charts
            bin_width = (max_val_m - min_val_m) / 10
            ' Places for x-values / Loss-values array
            places_array = (max_val_m * 1.2 - min_val_m * 0.7) / 0.1
            ReDim x_values(1 To places_array)
            ReDim Loss_values(1 To places_array)
            ' Determine first x_value based on smallest value from original sample
            step_value = min_val_m * 0.7
            ' Compute the loss function
            For i = 1 To places_array
                x_values(i) = step_value
                Loss_values(i) = constant_k * ((step_value - target) ^ 2)
                step_value = step_value + 0.1
            Next i
            ' Computing and creating the distribution curve
            ReDim freq_array(1 To 13)
            ReDim ctr_his_arr(1 To 13)
            step_value = min_val_m - bin_width
            For i = 1 To 13
                ctr_his_arr(i) = step_value
                step_value = step_value + bin_width * 0.5
                freq_array(i) = step_value
                step_value = step_value + bin_width * 0.5
            Next i
            ' Determine bin frequency for distribution
            freq_histo_orig = WorksheetFunction.Frequency(rng_orig_m, freq_array)
            freq_histo_boot = WorksheetFunction.Frequency(boot_val_arr, freq_array)
            Set Orig_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
            With Orig_Loss
                .ChartType = xlXYScatterSmoothNoMarkers
                .HasLegend = True
                .Legend.Position = xlLegendPositionTop
                .HasTitle = True
                .ChartTitle.Text = "Taguchi's Nominal-The-Best Loss Function"
                .Axes(xlCategory, xlPrimary).HasTitle = True
                .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
                .Axes(xlValue, xlPrimary).HasTitle = True
                .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
```

```
        .Axes(xlValue).MinimumScale = 0
        .Axes(xlCategory).MinimumScale = 0
        .Axes(xlCategory).MaximumScale = target * 2
        .SeriesCollection.NewSeries
        .SeriesCollection(1).name = "Target"
        .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
        .SeriesCollection(1).XValues = Array(target, target)
        .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
        .SeriesCollection.NewSeries
        .SeriesCollection(2).name = "LSL"
        .SeriesCollection(2).Values = Array(0, Loss_values(places_array))
        .SeriesCollection(2).XValues = Array(LSL, LSL)
        .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(31, 187, 61)
        .SeriesCollection.NewSeries
        .SeriesCollection(3).name = "USL"
        .SeriesCollection(3).Values = Array(0, Loss_values(places_array))
        .SeriesCollection(3).XValues = Array(USL, USL)
        .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(179, 39, 152)
        .SeriesCollection.NewSeries
        .SeriesCollection(4).name = "Loss"
        .SeriesCollection(4).XValues = x_values
        .SeriesCollection(4).Values = Loss_values
        .SeriesCollection(4).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
        .SeriesCollection.NewSeries
        .SeriesCollection(5).name = "Original Samples"
        .SeriesCollection(5).XValues = ctr_his_arr
        .SeriesCollection(5).Values = freq_histo_orig
        .SeriesCollection(5).AxisGroup = xlSecondary
        .Axes(xlValue, xlSecondary).MinimumScale = 0
        .SeriesCollection(5).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
        .Parent.Top = Worksheets("Simulation_Output").Cells(18, 3).Top
        .Parent.Left = Worksheets("Simulation_Output").Cells(18, 3 + dist).Left
        .Parent.Height = 224
        .Parent.Width = 340
    End With
    Set Boot_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
    With Boot_Loss
        .ChartType = xlXYScatterSmoothNoMarkers
        .HasLegend = True
        .Legend.Position = xlLegendPositionTop
        .HasTitle = True
        .ChartTitle.Text = "Taguchi's Nominal-The-Best Loss Function" + " Bootstrap (n = " + boot_iter_str + ")"
        .Axes(xlCategory, xlPrimary).HasTitle = True
        .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " "+ "[" + char_unit + "]"
        .Axes(xlValue, xlPrimary).HasTitle = True
        .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
        .Axes(xlValue).MinimumScale = 0
        .Axes(xlCategory).MinimumScale = 0
        .Axes(xlCategory).MaximumScale = target * 2
        .SeriesCollection.NewSeries
        .SeriesCollection(1).name = "Target"
        .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
        .SeriesCollection(1).XValues = Array(target, target)
        .SeriesCollection(1).AxisGroup = xlPrimary
        .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
        .SeriesCollection.NewSeries
        .SeriesCollection(2).name = "Bootstrap Data"
        .SeriesCollection(2).XValues = ctr_his_arr
        .SeriesCollection(2).Values = freq_histo_boot
        .SeriesCollection(2).AxisGroup = xlSecondary
        .Axes(xlValue, xlSecondary).MinimumScale = 0
        .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
        .SeriesCollection.NewSeries
        .SeriesCollection(3).name = "LSL"
        .SeriesCollection(3).Values = Array(0, Loss_values(places_array))
        .SeriesCollection(3).XValues = Array(LSL, LSL)
        .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(31, 187, 61)
```

```
    .SeriesCollection.NewSeries
    .SeriesCollection(4).name = "USL"
    .SeriesCollection(4).Values = Array(0, Loss_values(places_array))
    .SeriesCollection(4).XValues = Array(USL, USL)
    .SeriesCollection(4).Format.Line.ForeColor.RGB = RGB(179, 39, 152)
    .SeriesCollection.NewSeries
    .SeriesCollection(5).name = "Loss"
    .SeriesCollection(5).XValues = x_values
    .SeriesCollection(5).Values = Loss_values
    .SeriesCollection(5).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
    .Parent.Top = Worksheets("Simulation_Output").Cells(34, 3).Top
    .Parent.Left = Worksheets("Simulation_Output").Cells(34, 3 + dist).Left
    .Parent.Height = 224
    .Parent.Width = 340
End With
Call Histogram_means(mean, dist, boot_iter_str)
Call Histogram_variances(Variance, dist, boot_iter_str)
' Asymmetric cases NTB
Else
' Determining max and min value of original sample
max_val_m = WorksheetFunction.max(rng_orig_m)
min_val_m = WorksheetFunction.min(rng_orig_m)
' Computing bin width for histogram / charts
bin_width = (max_val_m - min_val_m) / 10
' Places for x-values / Loss-values array
places_array_max = (max_val_m * 1.2 - target) / 0.1
places_array_min = (target - min_val_m * 0.2) / 0.1
If places_array_max > places_array_min Then
    places_array = places_array_max
Else
    places_array = places_array_min
End If
' Determine first x_value based on smallest value from original sample
ReDim x_values(1 To places_array)
ReDim Loss_values(1 To places_array)
ReDim x_values_max(1 To places_array)
ReDim Loss_values_max(1 To places_array)
min_val_m = LSL * 0.5
' Compute the loss function for lower side of target
For i = 1 To places_array
    x_values(i) = min_val_m
    If min_val_m < target Then
        Loss_values(i) = cons_k_low * ((min_val_m - target) ^ 2)
    min_val_m = min_val_m + 0.1
    End If
Next i
' Compute the loss function for upper side of target
calc_target = target
For i = 1 To places_array
    x_values_max(i) = calc_target
    Loss_values_max(i) = cons_k_up * ((calc_target - target) ^ 2)
    calc_target = calc_target + 0.1
    If calc_target = max_val_m Then Exit For
Next i
' Computing and creating the distribution curve
ReDim freq_array(1 To 13)
ReDim ctr_his_arr(1 To 13)
min_val_m = WorksheetFunction.min(rng_orig_m)
step_value = min_val_m - bin_width
For i = 1 To 13
    ctr_his_arr(i) = step_value
    step_value = step_value + bin_width * 0.5
    freq_array(i) = step_value
    step_value = step_value + bin_width * 0.5
Next i
' Determine bin frequency for distribution
freq_histo_orig = WorksheetFunction.Frequency(rng_orig_m, freq_array)
```

```
freq_histo_boot = WorksheetFunction.Frequency(boot_val_arr, freq_array)
Set Orig_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
With Orig_Loss
.ChartType = xlXYScatterSmoothNoMarkers
.HasLegend = True
.Legend.Position = xlLegendPositionTop
.HasTitle = True
.ChartTitle.Text = "Taguchi's Nominal-The-Best Loss Function"
.Axes(xlCategory, xlPrimary).HasTitle = True
.Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
.Axes(xlValue, xlPrimary).HasTitle = True
.Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
.Axes(xlValue).MinimumScale = 0
.Axes(xlCategory).MinimumScale = 0
.Axes(xlCategory).MaximumScale = target * 2
.SeriesCollection.NewSeries
.SeriesCollection(1).name = "Target"
.SeriesCollection(1).Values = Array(0, Loss_values_max(places_array))
.SeriesCollection(1).XValues = Array(target, target)
.SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
.SeriesCollection.NewSeries
.SeriesCollection(2).name = "LSL"
.SeriesCollection(2).Values = Array(0, Loss_values_max(places_array))
.SeriesCollection(2).XValues = Array(LSL, LSL)
.SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(31, 187, 61)
.SeriesCollection.NewSeries
.SeriesCollection(3).name = "USL"
.SeriesCollection(3).Values = Array(0, Loss_values_max(places_array))
.SeriesCollection(3).XValues = Array(USL, USL)
.SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(179, 39, 152)
.SeriesCollection.NewSeries
.SeriesCollection(4).name = "Loss"
.SeriesCollection(4).XValues = x_values
.SeriesCollection(4).Values = Loss_values
.SeriesCollection(4).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
.SeriesCollection.NewSeries
.SeriesCollection(5).XValues = x_values_max
.SeriesCollection(5).Values = Loss_values_max
.SeriesCollection(5).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
.Legend.LegendEntries(5).Delete
.SeriesCollection.NewSeries
.SeriesCollection(6).name = "Input Data"
.SeriesCollection(6).XValues = ctr_his_arr
.SeriesCollection(6).Values = freq_histo_orig
.SeriesCollection(6).AxisGroup = xlSecondary
.Axes(xlValue, xlSecondary).MinimumScale = 0
.SeriesCollection(6).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
.Parent.Top = Worksheets("Simulation_Output").Cells(18, 3).Top
.Parent.Left = Worksheets("Simulation_Output").Cells(18, 3 + dist).Left
.Parent.Height = 224
.Parent.Width = 340
End With
' Loss Function chart with bootstrap distribution / drawn values for each bootstrap sample
Set Boot_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
With Boot_Loss
.ChartType = xlXYScatterSmoothNoMarkers
.HasLegend = True
.Legend.Position = xlLegendPositionTop
.HasTitle = True
.ChartTitle.Text = "Taguchi's Nominal-The-Best Loss Function" + " Bootstrap (n = " + boot_iter_str + ")"
.Axes(xlCategory, xlPrimary).HasTitle = True
.Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
.Axes(xlValue, xlPrimary).HasTitle = True
.Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
.Axes(xlValue).MinimumScale = 0
.Axes(xlCategory).MinimumScale = 0
.Axes(xlCategory).MaximumScale = target * 2
```

```
    .SeriesCollection.NewSeries
    .SeriesCollection(1).name = "Target"
    .SeriesCollection(1).Values = Array(0, Loss_values_max(places_array))
    .SeriesCollection(1).XValues = Array(target, target)
    .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
    .SeriesCollection.NewSeries
    .SeriesCollection(2).name = "LSL"
    .SeriesCollection(2).Values = Array(0, Loss_values_max(places_array))
    .SeriesCollection(2).XValues = Array(LSL, LSL)
    .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(31, 187, 61)
    .SeriesCollection.NewSeries
    .SeriesCollection(3).name = "USL"
    .SeriesCollection(3).Values = Array(0, Loss_values_max(places_array))
    .SeriesCollection(3).XValues = Array(USL, USL)
    .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(179, 39, 152)
    .SeriesCollection.NewSeries
    .SeriesCollection(4).name = "Loss"
    .SeriesCollection(4).XValues = x_values
    .SeriesCollection(4).Values = Loss_values
    .SeriesCollection(4).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
    .SeriesCollection.NewSeries
    .SeriesCollection(5).XValues = x_values_max
    .SeriesCollection(5).Values = Loss_values_max
    .SeriesCollection(5).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
    .Legend.LegendEntries(5).Delete
    .SeriesCollection.NewSeries
    .SeriesCollection(6).name = "Input Data"
    .SeriesCollection(6).XValues = ctr_his_arr
    .SeriesCollection(6).Values = freq_histo_boot
    .SeriesCollection(6).AxisGroup = xlSecondary
    .Axes(xlValue, xlSecondary).MinimumScale = 0
    .SeriesCollection(6).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
    .Parent.Top = Worksheets("Simulation_Output").Cells(34, 3).Top
    .Parent.Left = Worksheets("Simulation_Output").Cells(34, 3 + dist).Left
    .Parent.Height = 224
    .Parent.Width = 340
    End With
    Call Histogram_means_asym(Mean_upper, Mean_lower, dist, boot_iter_str)
    Call Histogram_var_asym(Variance_upper, Variance_lower, dist, boot_iter_str)
End If
Case "Smaller-the-better"
    ' Determining max and min value of original sample
    max_val_m = WorksheetFunction.max(rng_orig_m)
    min_val_m = WorksheetFunction.min(rng_orig_m)
    ' Computing bin width for histogram / charts
    bin_width = (max_val_m - min_val_m) / 10
    ' Places for x-values / Loss-values array
    places_array = (max_val_m * 1.5 - min_val_m * 0.7) / 0.1
    ReDim x_values(1 To places_array)
    ReDim Loss_values(1 To places_array)
    ' Determine first x_value based on smallest value from original sample
    step_value = 0
    ' Compute the loss function
    For i = 1 To places_array
        x_values(i) = step_value
        Loss_values(i) = constant_k * (step_value ^ 2)
        step_value = step_value + 0.1
    Next i
    ReDim freq_array(1 To 13)
    ReDim ctr_his_arr(1 To 13)
    ' Computing and creating the distribution curve
    step_value = min_val_m - bin_width
    For i = 1 To 13
        ctr_his_arr(i) = step_value
        step_value = step_value + bin_width * 0.5
        freq_array(i) = step_value
        step_value = step_value + bin_width * 0.5
```

```
Next i
' Determine bin frequency for distribution
freq_histo_orig = WorksheetFunction.Frequency(rng_orig_m, freq_array)
freq_histo_boot = WorksheetFunction.Frequency(boot_val_arr, freq_array)
Set Orig_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
With Orig_Loss
    .ChartType = xlXYScatterSmoothNoMarkers
    .HasLegend = True
    .Legend.Position = xlLegendPositionTop
    .HasTitle = True
    .ChartTitle.Text = "Taguchi's Smaller-the-better Loss Function"
    .Axes(xlCategory, xlPrimary).HasTitle = True
    .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
    .Axes(xlValue, xlPrimary).HasTitle = True
    .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
    .Axes(xlValue).MinimumScale = 0
    .Axes(xlCategory).MinimumScale = 0
    .Axes(xlCategory).MaximumScale = WorksheetFunction.max(x_values)
    .SeriesCollection.NewSeries
    .SeriesCollection(1).name = "Target"
    .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
    .SeriesCollection(1).XValues = Array(target, target)
    .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
    .SeriesCollection.NewSeries
    .SeriesCollection(2).name = "Loss"
    .SeriesCollection(2).XValues = x_values
    .SeriesCollection(2).Values = Loss_values
    .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
    .SeriesCollection.NewSeries
    .SeriesCollection(3).name = "Original Samples"
    .SeriesCollection(3).XValues = ctr_his_arr
    .SeriesCollection(3).Values = freq_histo_orig
    .SeriesCollection(3).AxisGroup = xlSecondary
    .Axes(xlValue, xlSecondary).MinimumScale = 0
    .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
    .Parent.Top = Worksheets("Simulation_Output").Cells(18, 3).Top
    .Parent.Left = Worksheets("Simulation_Output").Cells(18, 3 + dist).Left
    .Parent.Height = 224
    .Parent.Width = 340
End With
Set Boot_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
With Boot_Loss
    .ChartType = xlXYScatterSmoothNoMarkers
    .HasLegend = True
    .Legend.Position = xlLegendPositionTop
    .HasTitle = True
    .ChartTitle.Text = "Taguchi's Smaller-the-better Loss Function" + " Bootstrap (n = " + boot_iter_str + ")"
    .Axes(xlCategory, xlPrimary).HasTitle = True
    .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
    .Axes(xlValue, xlPrimary).HasTitle = True
    .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
    .Axes(xlValue).MinimumScale = 0
    .Axes(xlCategory).MinimumScale = 0
    .Axes(xlCategory).MaximumScale = WorksheetFunction.max(x_values)
    .SeriesCollection.NewSeries
    .SeriesCollection(1).name = "Target"
    .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
    .SeriesCollection(1).XValues = Array(target, target)
    .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
    .SeriesCollection.NewSeries
    .SeriesCollection(2).name = "Loss"
    .SeriesCollection(2).XValues = x_values
    .SeriesCollection(2).Values = Loss_values
    .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
    .SeriesCollection.NewSeries
    .SeriesCollection(3).name = "Bootstrap Data"
    .SeriesCollection(3).XValues = ctr_his_arr
```

```
            .SeriesCollection(3).Values = freq_histo_boot
            .SeriesCollection(3).AxisGroup = xlSecondary
            .Axes(xlValue, xlSecondary).MinimumScale = 0
            .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
            .Parent.Top = Worksheets("Simulation_Output").Cells(34, 3).Top
            .Parent.Left = Worksheets("Simulation_Output").Cells(34, 3 + dist).Left
            .Parent.Height = 224
            .Parent.Width = 340
        End With
        Call Histogram_means(mean, dist, boot_iter_str)
        Call Histogram_variances(Variance, dist, boot_iter_str)
    Case "Larger-the-better"
        ' Determining max and min value of original sample
        max_val_m = WorksheetFunction.max(rng_orig_m)
        min_val_m = WorksheetFunction.min(rng_orig_m)
        ' Computing bin width for histogram / charts
        bin_width = (max_val_m - min_val_m) / 10
        ' Places for x-values / Loss-values array
        places_array = (max_val_m * 1.8 - min_val_m * 0.7) / 0.1
        ReDim x_values(1 To places_array)
        ReDim Loss_values(1 To places_array)
        ' Determine first x_value based on smallest value from original sample
        step_value = 1.5
        ' Compute the loss function
        For i = 1 To places_array
            x_values(i) = step_value
            Loss_values(i) = constant_k / (step_value ^ 2)
            step_value = step_value + 0.1
        Next i
        ReDim freq_array(1 To 13)
        ReDim ctr_his_arr(1 To 13)
        ' Computing and creating the distribution curve
        step_value = min_val_m - bin_width
        For i = 1 To 13
            ctr_his_arr(i) = step_value
            step_value = step_value + bin_width * 0.5
            freq_array(i) = step_value
            step_value = step_value + bin_width * 0.5
        Next i
        ' Determine bin frequency for distribution
        freq_histo_orig = WorksheetFunction.Frequency(rng_orig_m, freq_array)
        freq_histo_boot = WorksheetFunction.Frequency(boot_val_arr, freq_array)
        Set Orig_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
        With Orig_Loss
            .ChartType = xlXYScatterSmoothNoMarkers
            .HasLegend = True
            .Legend.Position = xlLegendPositionTop
            .HasTitle = True
            .ChartTitle.Text = "Taguchi's Larger-the-better Loss Function"
            .Axes(xlCategory, xlPrimary).HasTitle = True
            .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
            .Axes(xlValue, xlPrimary).HasTitle = True
            .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
            .Axes(xlValue).MinimumScale = 0
            .Axes(xlCategory).MinimumScale = 0
            .Axes(xlCategory).MaximumScale = WorksheetFunction.max(x_values)
            .SeriesCollection.NewSeries
            .SeriesCollection(1).name = "Target"
            .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
            .SeriesCollection(1).XValues = Array(target, target)
            .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
            .SeriesCollection.NewSeries
            .SeriesCollection(2).name = "Loss"
            .SeriesCollection(2).XValues = x_values
            .SeriesCollection(2).Values = Loss_values
            .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
            .SeriesCollection.NewSeries
```

```
                .SeriesCollection(3).name = "Original Samples"
                .SeriesCollection(3).XValues = ctr_his_arr
                .SeriesCollection(3).Values = freq_histo_orig
                .SeriesCollection(3).AxisGroup = xlSecondary
                .Axes(xlValue, xlSecondary).MinimumScale = 0
                .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
                .Parent.Top = Worksheets("Simulation_Output").Cells(18, 3).Top
                .Parent.Left = Worksheets("Simulation_Output").Cells(18, 3 + dist).Left
                .Parent.Height = 224
                .Parent.Width = 340
            End With
            Set Boot_Loss = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
            With Boot_Loss
                .ChartType = xlXYScatterSmoothNoMarkers
                .HasLegend = True
                .Legend.Position = xlLegendPositionTop
                .HasTitle = True
                .ChartTitle.Text = "Taguchi's Larger-The-Better Loss Function" + " Bootstrap (n = " + boot_iter_str + ")"
                .Axes(xlCategory, xlPrimary).HasTitle = True
                .Axes(xlCategory, xlPrimary).AxisTitle.Text = char_name + " " + "[" + char_unit + "]"
                .Axes(xlValue, xlPrimary).HasTitle = True
                .Axes(xlValue, xlPrimary).AxisTitle.Text = "Loss" + " " + "[" + currency_unit + "]"
                .Axes(xlValue).MinimumScale = 0
                .Axes(xlCategory).MinimumScale = 0
                .Axes(xlCategory).MaximumScale = WorksheetFunction.max(x_values)
                .SeriesCollection.NewSeries
                .SeriesCollection(1).name = "Target"
                .SeriesCollection(1).Values = Array(0, Loss_values(places_array))
                .SeriesCollection(1).XValues = Array(target, target)
                .SeriesCollection(1).Format.Line.ForeColor.RGB = RGB(153, 153, 153)
                .SeriesCollection.NewSeries
                .SeriesCollection(2).name = "Loss"
                .SeriesCollection(2).XValues = x_values
                .SeriesCollection(2).Values = Loss_values
                .SeriesCollection(2).Format.Line.ForeColor.RGB = RGB(255, 0, 0)
                .SeriesCollection.NewSeries
                .SeriesCollection(3).name = "Bootstrap Data"
                .SeriesCollection(3).XValues = ctr_his_arr
                .SeriesCollection(3).Values = freq_histo_boot
                .SeriesCollection(3).AxisGroup = xlSecondary
                .Axes(xlValue, xlSecondary).MinimumScale = 0
                .SeriesCollection(3).Format.Line.ForeColor.RGB = RGB(204, 204, 255)
                .Parent.Top = Worksheets("Simulation_Output").Cells(34, 3).Top
                .Parent.Left = Worksheets("Simulation_Output").Cells(34, 3 + dist).Left
                .Parent.Height = 224
                .Parent.Width = 340
            End With
        End Select
    ' Computing total loss; computed depending on loss function type
    Select Case Worksheets("User_Input").Cells(7, 9 + m_col).Value
        Case "Nominal-the-best"
            If Worksheets("Computations").Cells(4, 2 + m_col).Value = "Yes" Then
            Total_loss_series = Total_loss_series + Worksheets("Simulation_Output").Cells(14, 4 + dist).Value
            Total_Gal_loss_series = Total_Gal_loss_series + Worksheets("Simulation_Output").Cells(14, 7 + dist).Value
            Else
            Total_loss_series = Total_loss_series + Worksheets("Simulation_Output").Cells(15, 4 + dist).Value +
Worksheets("Simulation_Output").Cells(15, 7 + dist).Value
            Total_Gal_loss_series = Total_Gal_loss_series + Worksheets("Simulation_Output").Cells(16, 4 + dist).Value +
Worksheets("Simulation_Output").Cells(16, 7 + dist).Value
            End If
        Case "Smaller-the-better"
            Total_loss_series = Total_loss_series + Worksheets("Simulation_Output").Cells(14, 4 + dist).Value
            Total_Gal_loss_series = Total_Gal_loss_series + Worksheets("Simulation_Output").Cells(14, 7 + dist).Value
        Case "Larger-the-better"
            Total_loss_series = Total_loss_series + Worksheets("Simulation_Output").Cells(14, 6 + dist).Value
            Total_Gal_loss_series = Total_Gal_loss_series + Worksheets("Simulation_Output").Cells(14, 6 + dist).Value
    End Select
```

```
    ' Printing total loss
    Worksheets("Simulation_Output").Range("A6").Value = "Total Loss"
    Worksheets("Simulation_Output").Range("A7").Value = Total_loss_series
    Worksheets("Simulation_Output").Range("A11").Value = "Total Galton Loss"
    Worksheets("Simulation_Output").Range("A12").Value = Total_Gal_loss_series
    ' Background coloring
    k_str = m_col
    If Worksheets("Simulation_Output").Cells(4, 3 + dist).Value = "Component" + " " + k_str Then
        With Worksheets("Simulation_Output")
            .Range(.Cells(3, 2 + dist), .Cells(90, 9 + dist)).Interior.Color = RGB(231, 230, 230)
            .Range(.Cells(4, 3 + dist), .Cells(16, 8 + dist)).Interior.Color = RGB(255, 192, 0)
        End With
    End If
    ' Place holder to print results onto sheet
dist = dist + 7
Next m_col
' Create input for Dropdown-list in Sensitivity_Analysis sheet
Worksheets("Sensitivity_Analysis").ComboBox1.Clear
For k = 1 To num_com
    Worksheets("Sensitivity_Analysis").ComboBox1.AddItem k
Next k
End Sub
' This sub creates histograms for the means(j) for Nominal-the-best and smaller-the-better cases, also used for displaying larger-the-better loss distribution
Sub Histogram_means(boostrap_statistic() As Double, dist As Integer, boot_iter_str As String)
Dim max As Double, min As Double, ave As Double, ave_st As String
Dim num_bins As Integer, bin_width As Double
Dim i As Integer, MyChart As Chart
Dim histo_array() As Double, center_histo_array() As Double, values_histo_array() As Variant
num_bins = Worksheets("User_Input").Range("F31").Value
' Determining max and min value of bootstrap statistic
max = WorksheetFunction.max(boostrap_statistic)
min = WorksheetFunction.min(boostrap_statistic)
ave = WorksheetFunction.Average(boostrap_statistic)
ave_st = Round(ave, 2)
' Computing bin width for histogram / charts
bin_width = (max - min) / num_bins
' Computing and creating the distribution curve
ReDim histo_array(1 To num_bins + 1)
ReDim center_histo_array(1 To num_bins + 1)
For i = 1 To num_bins + 1
    center_histo_array(i) = Round(min, 2)
    min = min + bin_width * 0.5
    histo_array(i) = Round(min, 2)
    min = min + bin_width * 0.5
Next i
' Determine bin frequency for distribution
values_histo_array = WorksheetFunction.Frequency(boostrap_statistic, histo_array)
Set MyChart = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
        With MyChart
            .ChartType = xlColumnClustered
            .HasTitle = True
            .ChartTitle.Text = "Grand-Mean = " + ave_st + " Bootstrap (n = " + boot_iter_str + ")"
            .Axes(xlCategory, xlPrimary).HasTitle = True
            .Axes(xlCategory, xlPrimary).AxisTitle.Text = "Bootstrap Means"
            .Axes(xlValue, xlPrimary).HasTitle = True
            .Axes(xlValue, xlPrimary).AxisTitle.Text = "Frequency"
            .SeriesCollection.NewSeries
            .SeriesCollection(1).Values = values_histo_array
            .SeriesCollection(1).XValues = center_histo_array
            .Parent.Top = Worksheets("Simulation_Output").Cells(50, 3).Top
            .Parent.Left = Worksheets("Simulation_Output").Cells(50, 3 + dist).Left
            .Parent.Height = 224
            .Parent.Width = 340
        End With
End Sub
' This sub creates histograms for the variance(j) for Nominal-the-best and smaller-the-better cases
```

```
Sub Histogram_variances(Variance() As Double, dist As Integer, boot_iter_str As String)
'Declaring of all variables
Dim max As Double, min As Double, ave As Double, ave_st As String
Dim num_bins As Integer, bin_width As Double
Dim i As Integer, MyChart As Chart
Dim histo_array() As Double, center_histo_array() As Double, values_histo_array() As Variant
num_bins = Worksheets("User_Input").Range("F31").Value
max = WorksheetFunction.max(Variance)
min = WorksheetFunction.min(Variance)
ave = WorksheetFunction.Average(Variance)
ave_st = Round(ave, 2)
bin_width = (max - min) / num_bins
ReDim histo_array(1 To num_bins + 1)
ReDim center_histo_array(1 To num_bins + 1)
For i = 1 To num_bins + 1
   center_histo_array(i) = Round(min, 2)
   min = min + bin_width * 0.5
   histo_array(i) = Round(min, 2)
   min = min + bin_width * 0.5
Next i
values_histo_array = WorksheetFunction.Frequency(Variance, histo_array)
Set MyChart = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
        With MyChart
            .ChartType = xlColumnClustered
            .HasTitle = True
            .ChartTitle.Text = "Grand-Variance = " + ave_st + " Bootstrap (n = " + boot_iter_str + ")"
            .Axes(xlCategory, xlPrimary).HasTitle = True
            .Axes(xlCategory, xlPrimary).AxisTitle.Text = "Bootstrap variances"
            .Axes(xlValue, xlPrimary).HasTitle = True
            .Axes(xlValue, xlPrimary).AxisTitle.Text = "Frequency"
            .SeriesCollection.NewSeries
            .SeriesCollection(1).Values = values_histo_array
            .SeriesCollection(1).XValues = center_histo_array
            .Parent.Top = Worksheets("Simulation_Output").Cells(66, 3).Top
            .Parent.Left = Worksheets("Simulation_Output").Cells(66, 3 + dist).Left
            .Parent.Height = 224
            .Parent.Width = 340
        End With
End Sub
' This sub creates histograms for the means(j) for Nominal-the-best - asymmetric cases
Sub Histogram_means_asym(boot_stat_up() As Double, boot_stat_low() As Double, dist As Integer, boot_iter_str As String)
Dim max As Double, min As Double
Dim ave_up As Double, ave_up_st As String
Dim ave_low As Double, ave_low_st As String
Dim num_bins As Integer, bin_width As Double
Dim i As Long, m As Long, MyChart As Chart
Dim histo_array() As Double, center_histo_array() As Double, values_histo_array() As Variant
Dim mean_up_low() As Double, boot_iter As Long, size_mean_up_low As Long
num_bins = Worksheets("User_Input").Range("F31").Value
boot_iter = Worksheets("User_Input").Range("F30").Value
size_mean_up_low = 2 * boot_iter
ReDim mean_up_low(1 To size_mean_up_low)
m = 1
For i = 1 To boot_iter
   mean_up_low(m) = boot_stat_up(i)
   m = m + 1
   mean_up_low(m) = boot_stat_low(i)
   m = m + 1
Next i
max = WorksheetFunction.max(mean_up_low)
min = WorksheetFunction.min(mean_up_low)
ave_up = WorksheetFunction.Average(boot_stat_up)
ave_low = WorksheetFunction.Average(boot_stat_low)
ave_up_st = Round(ave_up, 2)
ave_low_st = Round(ave_low, 2)
bin_width = (max - min) / num_bins
ReDim histo_array(1 To num_bins + 1)
```

```
ReDim center_histo_array(1 To num_bins + 1)
For i = 1 To num_bins + 1
   center_histo_array(i) = Round(min, 1)
   min = min + bin_width * 0.5
   histo_array(i) = Round(min, 1)
   min = min + bin_width * 0.5
Next i
values_histo_array = WorksheetFunction.Frequency(mean_up_low, histo_array)
Set MyChart = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
         With MyChart
            .ChartType = xlColumnClustered
            .HasTitle = True
            .ChartTitle.Text = "Mean Upper Side = " + ave_up_st + " /" + "                    Mean Lower Side = " + ave_low_st
            .Axes(xlCategory, xlPrimary).HasTitle = True
            .Axes(xlCategory, xlPrimary).AxisTitle.Text = "Bootstrap Means"
            .Axes(xlValue, xlPrimary).HasTitle = True
            .Axes(xlValue, xlPrimary).AxisTitle.Text = "Frequency"
            .SeriesCollection.NewSeries
            .SeriesCollection(1).Values = values_histo_array
            .SeriesCollection(1).XValues = center_histo_array
            .Parent.Top = Worksheets("Simulation_Output").Cells(50, 3).Top
            .Parent.Left = Worksheets("Simulation_Output").Cells(50, 3 + dist).Left
            .Parent.Height = 224
            .Parent.Width = 340
         End With
End Sub
' This sub creates histograms for the variances(j) for Nominal-the-best - asymmetric cases
Sub Histogram_var_asym(boot_stat_up() As Double, boot_stat_low() As Double, dist As Integer, boot_iter_str As String)
Dim max As Double, min As Double
Dim ave_up As Double, ave_up_st As String
Dim ave_low As Double, ave_low_st As String
Dim num_bins As Integer, bin_width As Double
Dim i As Long, m As Long, MyChart As Chart
Dim histo_array() As Double, center_histo_array() As Double, values_histo_array() As Variant
Dim var_up_low() As Double, boot_iter As Long, size_var_up_low As Long
num_bins = Worksheets("User_Input").Range("F31").Value
boot_iter = Worksheets("User_Input").Range("F30").Value
size_var_up_low = 2 * boot_iter
ReDim var_up_low(1 To size_var_up_low)
m = 1
For i = 1 To boot_iter
   var_up_low(m) = boot_stat_up(i)
   m = m + 1
   var_up_low(m) = boot_stat_low(i)
   m = m + 1
Next i
max = WorksheetFunction.max(var_up_low)
min = WorksheetFunction.min(var_up_low)
ave_up = WorksheetFunction.Average(boot_stat_up)
ave_low = WorksheetFunction.Average(boot_stat_low)
ave_up_st = Round(ave_up, 2)
ave_low_st = Round(ave_low, 2)
bin_width = (max - min) / num_bins
ReDim histo_array(1 To num_bins + 1)
ReDim center_histo_array(1 To num_bins + 1)
For i = 1 To num_bins + 1
   center_histo_array(i) = Round(min, 1)
   min = min + bin_width * 0.5
   histo_array(i) = Round(min, 1)
   min = min + bin_width * 0.5
Next i
values_histo_array = WorksheetFunction.Frequency(var_up_low, histo_array)
Set MyChart = Worksheets("Simulation_Output").Shapes.AddChart2.Chart
         With MyChart
            .ChartType = xlColumnClustered
            .HasTitle = True
```

```
            .ChartTitle.Text = "Variance Upper Side = " + ave_up_st + " /" + "                    Variance Lower Side = " +
ave_low_st
            .Axes(xlCategory, xlPrimary).HasTitle = True
            .Axes(xlCategory, xlPrimary).AxisTitle.Text = "Bootstrap Variances"
            .Axes(xlValue, xlPrimary).HasTitle = True
            .Axes(xlValue, xlPrimary).AxisTitle.Text = "Frequency"
            .SeriesCollection.NewSeries
            .SeriesCollection(1).Values = values_histo_array
            .SeriesCollection(1).XValues = center_histo_array
            .Parent.Top = Worksheets("Simulation_Output").Cells(66, 3).Top
            .Parent.Left = Worksheets("Simulation_Output").Cells(66, 3 + dist).Left
            .Parent.Height = 224
            .Parent.Width = 340
        End With
End Sub
Private Sub Simulation_Design_Click()
Dim k As Integer, k_str As String
Dim num_com As Integer
num_com = Worksheets("User_Input").Range("F21").Value
' Clears already existing table from former simulation process
Worksheets("User_Input").Range(Cells(5, 9), Cells(500, 21)).Clear
'Creating Table for Data Input
With Worksheets("User_Input")
    .Range("I5:I13").Font.size = 14
    .Range("I5:I13").HorizontalAlignment = xlCenter
    .Range("I5:I13").VerticalAlignment = xlCenter
    .Range("I5").Value = "Variables"
    .Range("I6").Value = "Name"
    .Range("I7").Value = "Loss Function Type"
    .Range("I8").Value = "Target"
    .Range("I9").Value = "Upper Spec Limit (USL)"
    .Range("I10").Value = "Lower Spec Limit (LSL)"
    .Range("I11").Value = "Loss at USL / Target"
    .Range("I12").Value = "Loss at LSL"
    .Range("I13").Value = "Enter your values here"
    'Creating Excel Table Style
    .ListObjects.Add(xlSrcRange, Range(Cells(5, 9), Cells(13, 9)), , xlYes, , "TableStyleMedium17").name =
"Simulation_Design_Table"
    .ListObjects("Simulation_Design_Table").Range.AutoFilter
End With
'Determining number of columns for investigated components
For k = 1 To num_com
    k_str = k
    Worksheets("User_Input").Cells(5, 9 + k).Value = "Component" + " " + k_str
    Worksheets("User_Input").Cells(5, 9 + k).ColumnWidth = 22.3
    ' Creating Drop-Down-Menu for user to select loss function
    If Not IsEmpty(Worksheets("User_Input").Cells(5, 9 + k).Value) Then
        Cells(7, 9 + k).Select
        With Selection.Validation
            .Add Type:=xlValidateList, AlertStyle:=xlValidAlertStop, Operator:= _
            xlBetween, Formula1:="=$G$22:$G$24"
            .IgnoreBlank = True
            .InCellDropdown = True
            .InputTitle = ""
            .ErrorTitle = ""
            .InputMessage = ""
            .ErrorMessage = ""
            .ShowInput = True
            .ShowError = True
    End With
    End If
Next k
End Sub
```

Appendix B

Table 16. Sensitivity analysis of the average loss per unit for Taguchi's *nominal-the-best* quality loss function for shifted mean for a simplified biomass supply chain with a cost constant $k = 2$ \$/%².

Average Loss per unit in \$ in terms of shifted mean per 0.5 sigma				
Component Sigma	Harvest / Collection	Transport	Drying	Densification
-6	268.01	212.89	217.33	132.43
-5.5	226.31	179.45	184.11	110.18
-5	188.24	148.95	153.72	90.06
-4.5	153.80	121.39	126.16	72.05
-4	123.00	96.78	101.44	56.16
-3.5	95.82	75.12	79.56	42.39
-3	72.28	56.40	60.51	30.74
-2.5	52.36	40.62	44.29	21.20
-2	36.08	27.79	30.91	13.78
-1.5	23.43	17.91	20.36	8.48
-1	14.41	10.97	12.65	5.30
-0.5	9.02	6.97	7.78	4.24
0	7.26	5.92	5.74	5.29
0.5	9.14	7.81	6.53	8.46
1	14.64	12.65	10.16	13.75
1.5	23.77	20.44	16.63	21.16
2	36.54	31.17	25.93	30.69
2.5	52.94	44.84	38.06	42.33
3	72.96	61.46	53.03	56.09
3.5	96.62	81.03	70.84	71.97
4	123.91	103.54	91.47	89.97
4.5	154.83	128.99	114.95	110.09
5	189.38	157.39	141.26	132.32
5.5	227.57	188.74	170.40	156.67
6	269.38	223.03	202.38	183.14

Table 17. Sensitivity analysis of the average loss per unit for Taguchi's *nominal-the-best* quality loss function in case of variance for a simplified biomass supply chain for cost constant $k = 2$ \$/%².

Average Loss per unit in \$ in terms of changing variance per 0.5 sigma				
Component Sigma	Harvest / Collection	Transport	Drying	Densification
-6	1.63	2.05	1.55	0.15
-5.5	2.10	2.37	1.90	0.58
-5	2.57	2.69	2.24	1.01
-4.5	3.04	3.02	2.59	1.44
-4	3.51	3.34	2.94	1.87
-3.5	3.98	3.66	3.29	2.29
-3	4.45	3.98	3.64	2.72
-2.5	4.91	4.31	3.99	3.15
-2	5.38	4.63	4.34	3.58
-1.5	5.85	4.95	4.69	4.01
-1	6.32	5.27	5.04	4.43
-0.5	6.79	5.60	5.39	4.86
0	7.26	5.92	5.74	5.29
0.5	7.73	6.24	6.09	5.72
1	8.20	6.57	6.44	6.15
1.5	8.67	6.89	6.79	6.57
2	9.14	7.21	7.14	7.00
2.5	9.61	7.53	7.49	7.43
3	10.08	7.86	7.83	7.86
3.5	10.55	8.18	8.18	8.29
4	11.02	8.50	8.53	8.71
4.5	11.49	8.82	8.88	9.14
5	11.96	9.15	9.23	9.57
5.5	12.43	9.47	9.58	10.00
6	12.90	9.79	9.93	10.43

Appendix C

Continuous Improvement Handbook for the Sustainable Bio-Based Industries

Christoph Metzner
The University of Tennessee
Center for Renewable Carbon

Knoxville, December 2018

This handbook was developed within the scope of the master thesis "Strategies for continuous improvement and competitiveness for the sustainable bio-based industries".

Special gratitude is owned to Dr. Timothy M. Young for his ongoing advisory throughout the creation of the thesis and this handbook.

Table of Content

List of Tables

List of Figures

Purpose and Target Group

The following handbook was developed to introduce practitioners of the sustainable bio-based industries into continuous improvement concepts. Concepts which help to identify key process variables and detect main sources of their underlying variations. Furthermore, to show the importance and benefits of visualizing and quantifying variation through use of statistical methodologies.

Background

Quality management is crucial for manufacturing organizations not only for profitability competitiveness but for business survival. Many consider the American quality revolution of the 1980s to have been a response to an American quality crisis. The conditions of the business quality crisis of the 1980s are very similar to the current economic crisis faced by U.S. businesses and especially the sustainable bio-based industries. The current economic crisis of unprecedented oil prices and a severe decline in housing starts with constrained capital markets and had some forest products companies renewing their emphasis on variation reduction, cost savings, and quality improvement.

Why Variation in Manufacturing is Important

Variation among all components of a production is a vital factor for determining the performance and success of your company. Generally, there are five sources for process variation involved in a manufacturing process which are the following: Raw materials, equipment, human actions, environment, and production methodology. Understanding that raw material quality is subject to natural variation is essential for a successful performance. In addition, products from several suppliers could be different, *i.e., check the variation of key attributes of the raw material to identify the best supplier*. Equipment varies in complexity and quality based on process and producer. Highly complex machinery could introduce more variation than simpler ones. Many factors could affect humans' actions during the manufacturing process; for example, even though two different employees were trained in the same way for operating an equipment they could have slightly different operating decision-making which is another important source of variation. Even for the same employee, he/she could also act differently between today and the next day, which also causes variation. For example, processing of sawn timber certain environmental settings such as the temperature or moisture content in the air are required. Any changes to these settings could lead to additional unnecessary variation in the final product. Lastly, manufacturing a product consists out of a series of steps which creates an optimal output; slight changes of this very order introduces variation. Overall, too many factors could cause variation in the manufacturing process. Hence, the ability to detect and reduce variation in manufacturing became a vital role for continuous improvement and competitiveness in the sustainable bio-based industries.

Quantifying Variation of Key Production Characteristics

The sustainable bio-based industries face serious challenges regarding higher raw material prices and quality variation and increased competition through globalization. Thus, making it inevitable that we think about adapting the current state of producing products to a modernized and more data analytics driven manufacturing direction. Efficient material input through visualizing and quantifying material variation is a key way in overcoming those challenges, *i.e., make your production more visible, efficient, and lean.* Therefore, a simulation tool was developed to offer practitioners a method to quantify variation of product attributes regarding supply chain or production related operations in monetary terms. The ultimate goal of this simulation tool is to identify the component of your system which induces the most loss due to variation. Identifying this component will help you to make better informed decisions. Furthermore, the simulation computes the loss for a series of components and even provides figures which emphasizes the impacts of the components on each other. The simulation tool can be downloaded under the link www.spc4lean.com.

Continuous Improvement Strategies

Continuous improvement (Kaizen, jap.) is a company-wide applied philosophy and describes tools and concepts to enhance the performance of the enterprise [1]. Furthermore, continuous improvement is a never-ending process of the current state and describes incremental improvement by little steps through participation of all entities and people in the company [2].

In contrast to the reactionary traditional quality control, continuous improvement philosophies are proactive with focus on prevention and early detection of problems. Decision making is based on defendable information from acceptable statistical methodology, while opinions are discounted. Continuous improvement strategies such as Statistical Process Control (SPC) enables visualization and quantification of variation with control charts. SPC is a general term that applies practical statistical methods to manufacturing process with the goal of variation reduction. Variation reduction of key process variables and product attributes leads to target reduction and cost savings.

Fundamental Requirements for Optimization

As already mentioned continuous improvement is not a tool which you just implement once and then hope it will work. Continuous improvement is a philosophy, a culture everybody in the company from the top management to the worker must adapt to, more must live the philosophy. Thus, change is dependent on the culture of the company, i.e., if the present culture of the work force and management will not accept data driven decision-making, and the use of statistical methods to diagnose sources of variation that lead to avoidable costs, do not dedicate resources to continuous improvement at this type of manufacturing facility.

Provide an environment to store your sensor data in a reliable data warehouse system and a secure electronic database of your destructive test results [3]. Apply modern data fusion methodologies to analyze your data real-time. Maintain a high-quality database, because this is necessary to initiate continuous improvement [3].

Visualize your Production – Process Flow Chart

Start your continuous improvement journey by identifying all steps of your production or process by creating a process flow chart. A process flow chart helps to show the logical sequence of all activities related towards manufacturing the product. Analyze and identify in a group of technicians and engineers non-value adding activities, e.g. unnecessary movement or storage of the product / materials and eliminate them [4].

Typically, standardized symbols are used to emphasize certain actions. *For example, see the process flow chart of the full-cell and modified full-cell pressure treating process for treated lumber (Figure 1).*

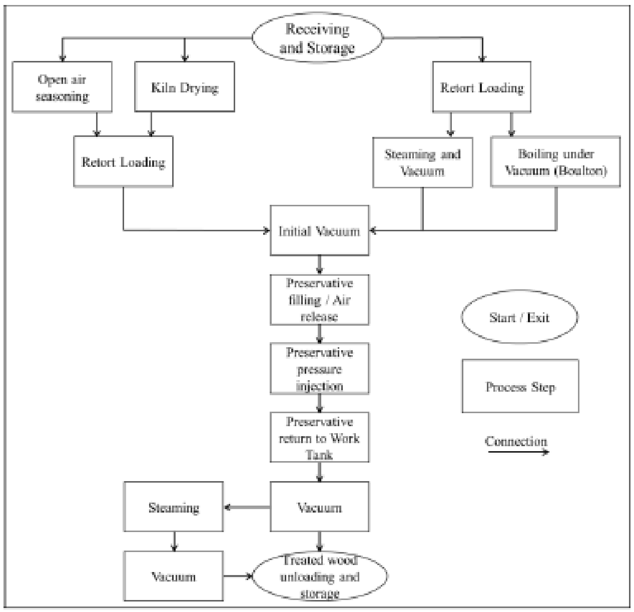


Figure 1. Example Process Flow Chart: Full-cell and modified full-cell pressure treating process for treated lumber [5].

Connection to the Customer -

Connect Process Variables with Product Attributes

There are at least two clienteles interested in your product, your customer and you as the manufacturer. However, you are not necessarily sharing the same interests in the product as the customer. Typically, the customer is interested in the attributes (*i.e.*, quality characteristics) of the finished product, such as *mechanical properties*, *thickness swell*, or *protection ability against organisms of wood preservatives*. Instead, you are interested in key process variables of your manufacturing system, like *moisture content, line speed, or chamber pressure*. Therefore, as a top priority you should connect the crucial attributes favored by the customer with the so-called *'critical few'* process variables (Figure 2) [3]. These process variables can be directly linked to the product attributes.

Take advantage of market analysis to gain information about or explicitly ask your customer about the desired product attributes. Furthermore, use statistical analysis and methods such as *Design of Experiments* or *Data Mining* to identify the *'critical few'* process variables [3]. Please be aware that many statistical methodologies only provide accurate answers for processes in the state of control; *i.e., constant variance*. However, it is very crucial for the continuous improvement process to just focus on the *'critical few'* process variables. Don't waste time and effort by charting and analyzing every variable, just because it exists.

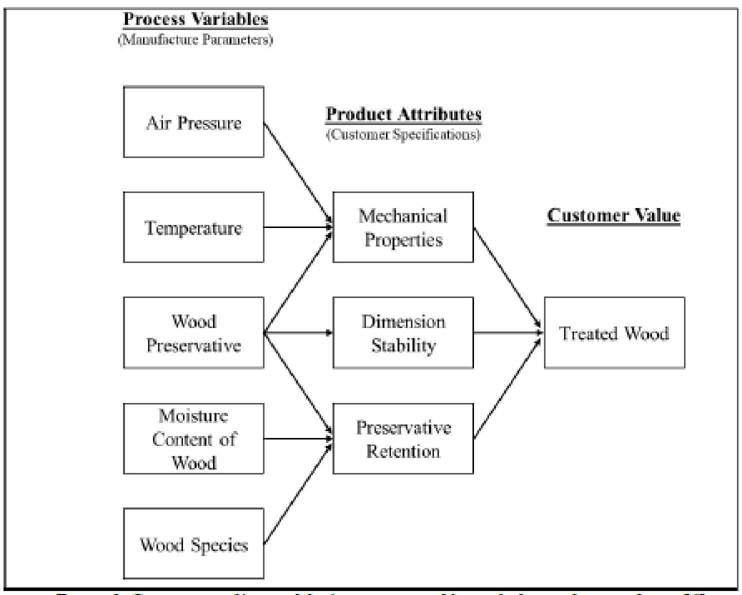


Figure 2. Connection of 'critical few' process variables with the product attributes [6].

Describe your Data

It is impossible to measure everything in raw material manufacturing processes, *e.g., the thickness of every pole at every position*. Instead, we make a few measurements (a "sample") and use these data to make inferences about the overall process (the "population"). Taken by themselves, these sample numbers have little meaning. However, by applying the proper statistical techniques the data can provide valuable information about your system and its variation.

Descriptive Statistics

Unlike inferential statistics which tries to draw inferences from a sample data about the underlying population descriptive statistics just describe sample data. Key statistics to describe sample data are introduced on the following page.

Average or Mean is the "center of the data set" $(\overline{X})$.

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

The **Median** of a set of measurements is the middle value when the measurements are arranged from smallest to largest. Also known as the 50th percentile.

The **Range** of a data set is the distance from the minimum to the maximum value.

$$Range = Maximum - Minimum$$

The **Standard Deviation** estimates the variability in the same unit of measure.

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

The **Variance** measures how far the values from a data set are spread from their average. The variance is the square of the standard deviation; $s = s^2$.

The **Coefficient of Variation** (CV) Is a scaled measure of dispersion, which is the standard deviation divided by the mean (multiplied by one hundred percent). Helpful when comparing dispersion statistics across sets of data with varying scales or measure and means, e.g., product types, etc.

$$CV = \frac{s}{\overline{X}} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}}{\frac{\sum_{i=1}^{n} x_i}{n}} \times 100\%$$

For example, the following moisture contents were measured for dried lumber used in the wood treatment production.

26.8%	*26.6%*	*27.2%*	*24.4%*	*25%*	*27.4%*	*23.9%*

Average:

$$\overline{X} = \frac{26.8 + 26.6 + 27.2 + 24.4 + 25 + 27.4 + 23.9}{7} = 25.9\%$$

Median:

23.9%	*24.4%*	*25%*	$\boxed{26.6\%}$	*26.8%*	*27.2%*	*27.4%*

Range: *27.4% - 23.9% = 3.5%*

Standard Deviation:

$$s = \sqrt{\frac{(26.8-25.9)^2 + (26.6-25.9)^2 + (27.2-25.9)^2 + (24.4-25.9)^2 + (25-25.9)^2 + (27.4-25.9)^2 + (23.9-25.9)^2}{7-1}}$$

s = 1.43 %

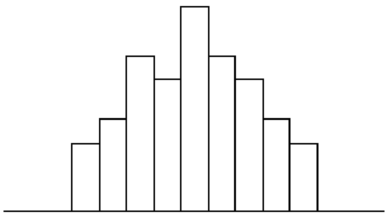
Variance: $s^2 = 2.05$

Coefficient of Variation:

$$CV = \frac{1.43}{25.9} \times 100\% = 5.5\%$$

How is your Data Distributed?

Generally, before using statistical methodologies analyze your data. Which type of distribution does your data follow (*e.g.*, normal, lognormal, Weibull, etc.)? Is your data set symmetric? Is your data skewed (e.g., left or right)? Is your data kurtosis (Figure 3)? Do you have outliers? Think about how outliers were created, i.e. was the measurement correct, did something happen? Your outliers affect your average; you might want to use the median instead. Do not manipulate or clean your data without justifying it!!!



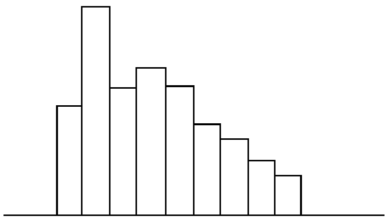


Figure 3. Histograms for data with symmetric distribution (left) and right-skewed distribution (right).

The Normal Distribution

Samples vary from one to the next, even if the process is not changing, for example some sawn timber boards are thicker than others and by chance you will measure thicker and thinner boards if you keep measuring. Often, if you make a lot of measurements and plot them in a histogram, you will get a picture like the one shown below (Figure 4). This picture shows a *Normal* or *bell-shaped* distribution. Calling it *normal*' does not mean that there are *abnormal* data, or that *Normal* is good – it is just a name! But *Normal Distributions* are often found in repeated sampling (i.e. Normal is normal), and statisticians have good rules for describing their properties, whether the sampling is of lumber thickness, moisture content or IB strength. If the distribution is Normal, it has predictable properties.

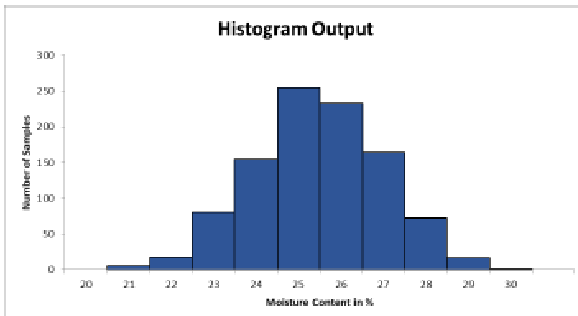


Figure 4. Typical histogram output for moisture content of lumber.

For data with a *Normal Distribution (symmetric)* the average will be in the middle; and very similar to the median. Furthermore, the average ± the standard deviation will cover the range of 68.2% of the data, ± 2 standard deviation 95% of the data, and ± 3 standard deviation 99% (Figure 5). The properties form the basis for control charting which will be discussed in the next chapter.

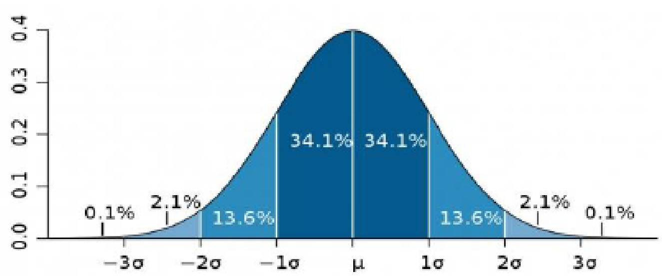


Figure 5. Percentage dispersion for a normally distributed data set in terms of standard deviation[1].

[1] http://news.mit.edu/2012/explained-sigma-0209

Variation – the Obstacle of Continuous Improvement

Variation the key obstacle for a (un-)successful performance of your production. As you know your production or process consists out of many different components, *e.g., machine settings, raw material quality, or machine parts*. Furthermore, every component will introduce variation in your process. Thus, observing and visualizing your variation is the first step in continuous improvement. This can be achieved by utilizing Statistical Process Control (SPC) tools, such as control charts. The goal of continuous improvement is the reduction of variation in your process. However, it is crucial to distinguish variation between common-cause or natural-cause variation and special-cause variation (Figure 6).

Common-cause variation comes from the system, *e.g., variation within a machine, variation between the same kind of machines (e.g., drum debarker), or variation between operators, etc.* Specifically, in the bio-based industries, natural variation mainly comes from the raw material such as moisture content, material thickness, and so on.

Special-cause variation results from an assignable cause, *e.g., machine stop, flaker blade damage, platen damage, shift-change, etc.*. Both natural and special cause variation represents a cost to any wood product manufacture. Importantly, most scholars agree that a process can only be improved after identifying, investigating, and eliminating all sources of special-cause variation [7].

Natural variation and special cause variation are accurately quantified by the use of Shewhart control charts. Implement control charts only for the key product attributes linked to the 'critical few' process variables to quantify natural variation and special cause variation. Control charts are an early detection tool for preventing problems and reducing scrap material. Do not use control charts for every variable, only those which impact your product the most. Train key management and operation personnel in the use and understanding of control charts.

Visualize your Variation – Control Charts

Variation during the processing time can be observed through control charts, specifically Shewhart control charts (Figure 6). In the control charts, x-axis generally represent processing time or sample and y-axis mostly denotes the individual measurements of a variable in the processing or represent ranges of variation among subgroups even the averages of subgroups. The main purpose of control charts is to show the performance of a process and how this process is impacted by changes. Overall, control charts are an essential tool for quality control and they can help detect whether the variation is a natural or a special cause variation (Figure 7). For some processes the application of multivariate control charts to visualize and monitor variation is more suitable.

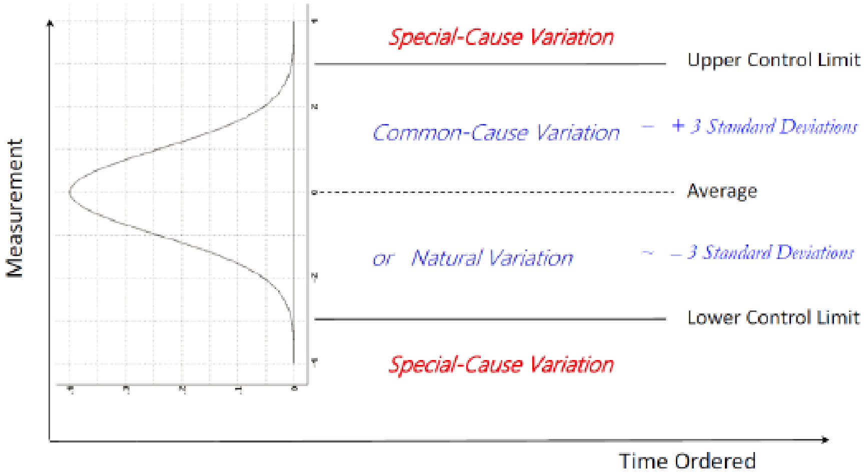


Figure 6. Context of control charts with common and special-cause variation for normally distributed data [8].

Control charts have four key features:

1) Data points are either averages of subgroup measurements or individual measurements plotted on the x/y axis and joined by a line. Time is always on the x-axis.

2) The Average or Center Line is the average or mean of the data points and is drawn across the middle section of the graph, usually as a heavy or solid line.

3) The Upper Control Limit (UCL) is drawn above the centerline and often annotated as "UCL". This is often called the "+ 3 sigma" line.

4) The Lower Control Limit (LCL) is drawn below the centerline and often annotated as "LCL". This is called the "- 3 sigma" line.

Specification limits are clear requirements which the material or product must satisfy. These limits stem from costumers or associations, such as American Wood Protection Association (AWPA).

Instead, control limits represent ± 3 standard deviation and are computed from the data.

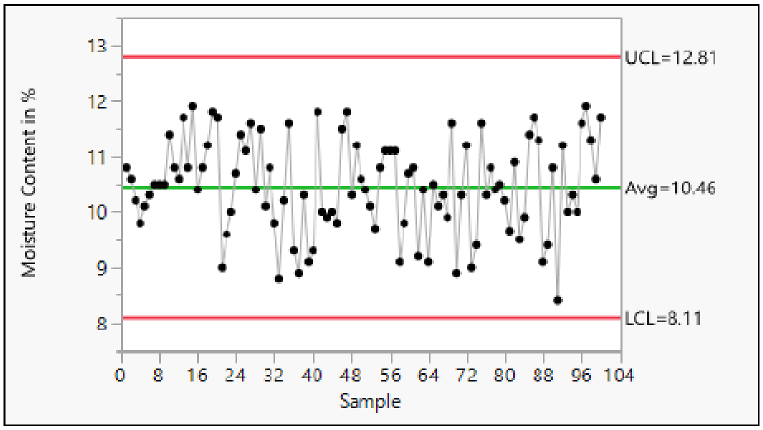


Figure 7: Control chart: X-individual chart.

Originally, four control run rules were introduced by Western Electric Company [9] and later updated to eight by Lloyd S. Nelson [10] to detect special-cause variation in control charts, which are the following:

1. One point is more than three standard deviations from the mean, *i.e., outlier indicates a process out of control.*
2. Nine (or more) points in a row are on the same side of the mean, *i.e., indicates a shift in the mean.*
3. Six (or more) points in a row are continually increasing (or decreasing), *i.e., indicates a trend.*
4. Fourteen (or more) points in a row alternate in direction, increasing then decreasing, *i.e., indicate at least two different data sets.*
5. Two (or three) out of three points in a row are more than two standard deviations from the mean in the same, *i.e., indicates a shift in the mean.*
6. Four (of five) out of five points in a row are more than one standard deviation from the mean in the same direction, *i.e., indicates a shift in the mean.*
7. Fifteen points in a row are all within one standard deviation of the mean on either side of the mean, *i.e., a higher variation would be expected.*
8. Eight points in a row exist, but none within one standard deviation of the mean, and the points are in both directions from the mean, *i.e., indicate at least two different data sets.*

Shewhart distinguished between control charts for measurement data and attribute data (Table 1). Measurement data come from continuous measurements and are considered a real number, *e.g., heights, densities, moisture content, physical dimensions, etc*. Attribute data are integers and are data, such as number of rejects, blemishes, etc.

Table 1 provides an overview of common control charts for measurement and attribute data.

Table 1. Typical univariate control charts for measurement and attribute data [7, 11].

Measurement Data				
Control Chart Type		Central Line	Control Limits	Purpose and when to use
Subgroup n = 1	X-Individual	$CL_X = \bar{X}$	$UCL_X = \bar{X} + 2.660\,\overline{mR}$ $LCL_X = \bar{X} - 2.660\,\overline{mR}$	Assessment of long-and short-term process variation – periodically collected data (organization of data in rational manner)
	Moving Range	$CL_R = \overline{mR}$	$UCL_R = 3.268\,\overline{mR}$	Assessment of stability of short-term process variation – slowly changing process
Subgroup n > 1	X-bar	$CL_{\bar{X}} = \bar{\bar{X}}$	$UCL_{\bar{X}} = \bar{\bar{X}} + A_2\,\bar{R}$ $LCL_{\bar{X}} = \bar{\bar{X}} - A_2\,\bar{R}$	Assessment of stability of the location of the process relative to its target – historical summary and organization of data into rational subgroups
	Range	$CL_R = \bar{R}$	$UCL_R = D_4\bar{R}$ $LCL_R = D_3\bar{R}$	Assessment of stability of the process variation within and between subgroups – historical summary and organization of data into rational subgroups
Attribute Data				
Binomial data	np chart	$CL_{np} = n\bar{p}$	$UCL_{np} = n\bar{p} + 3\sqrt{n\bar{p}(1-\bar{p})}$ $LCL_{np} = n\bar{p} - 3\sqrt{n\bar{p}(1-\bar{p})}$	n constant – all samples have the same sized areas of opportunity – counts bad and good samples
	p chart	$CL_p = \bar{p}$	$UCL_p = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$ $LCL_p = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$	n variable – Areas of opportunity changes from sample to sample – counts bad and good samples
Poisson data	c chart	$CL_c = \bar{c}$	$UCL_c = \bar{c} + 3\sqrt{\bar{c}}$ $LCL_c = \bar{c} - 3\sqrt{\bar{c}}$	a constant – all samples have the same sized areas of opportunity – used to count bad samples in complex products
	u chart	$CL_u = \bar{u}$	$UCL_u = \bar{u} + 3\sqrt{\frac{\bar{u}}{a_i}}$ $LCL_u = \bar{u} - 3\sqrt{\frac{\bar{u}}{a_i}}$	a variable – Areas of opportunity changes from sample to sample – used to count bad samples in complex products

For example, you want to analyze the moisture content of one batch kiln dried lumber (Table 2).

Long-term changes can be shown with the X-Individual control chart and short-term changes with the moving range control chart.

Table 2. Control Chart Example.

Sample	*Moisture Content [%]*	*Moving Range*
1	8.7	
2	9.7	$mR_1 = 9.7 - 8.7 = 1$
3	11.5	1.8
4	10	1.5
5	10.4	0.4
6	10.4	0
7	9.1	1.3
8	11	1.9
9	12	1
10	10.1	1.9
11	11	0.9
12	9.8	10.2
13	10.4	0.6
14	10.7	0.3
15	9.8	0.9

The average moisture content is: $\bar{x} = 10.31\%$

The average moving range is: $\overline{mR} = 1.05\%$

X-Individual Control Chart (Figure 8)
Upper Control Limit is: $UCL_x = \bar{X} + 2.660\,\overline{mR} = 10.31\% + 2.660 * 1.05\,\% = 13.10\,\%$
Lower Control Limit is: $LCL_x = \bar{X} - 2.660\,\overline{mR} = 10.31\% - 2.660 * 1.05\,\% = 7.52\%$
Natural variation of the process is: 13.10% - 7.52 % = 5.58%

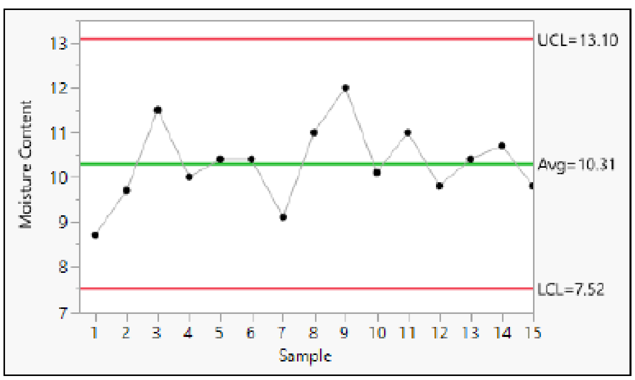


Figure 8. Example: X-Individual control chart for moisture content of kiln dried lumber.

Moving Range Control Chart (Figure 9)

Upper control limit: $UCL_R = 3.268\,\overline{mR}$ *= 3.268 * 1.05% = 3.430%*

A lower control limit for the moving range does not exist.

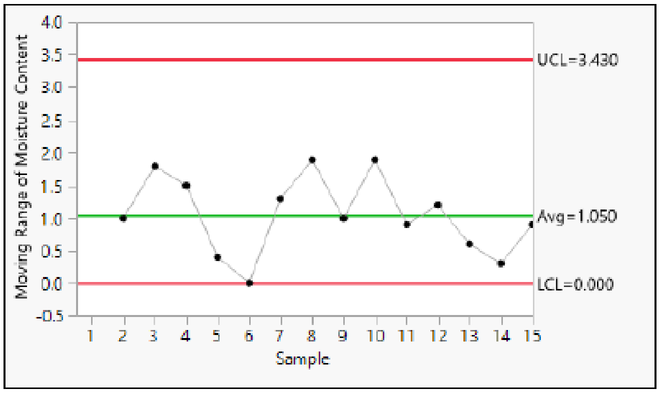


Figure 9. Example: Moving Range control chart for moisture content of kiln dried lumber.

As you can see whether for the X-Individual nor for the moving range chart values are out-of-spec nor special cause variation.

For control charts with subgroups specific constants are necessary, which can be found in the appendix.

Determine your Cost of Variation – The Taguchi Loss Function

Determine the financial cost caused by variation of your process variables or product attributes. This allows you to provide information and a greater foundation for managerial decision-making, as well as stronger incentives to act.

Traditionally, product quality is seen as the conformance to certain specification limits. Products within specifications are treated as equally good, while products outside specifications are treated as equally bad in terms of financial loss (Figure 10, left) [12]. That would mean, products on target are equally good in the sense of quality as products just meeting specification limits. But this is not the case. A product just within specification limits would rather be very similar towards a product just not meeting specification limits.

For example, assume the case of dried lumber right before the treatment process with wood preservatives in a vacuum chamber. The target is 25% with an upper specification limit of 28% and a lower specification of 22%. Furthermore, the workers know that the best quality for treated wood stems from lumber with an initial moisture content of 25%. In this sense, lumber with 28% and 29% moisture content would rather create the same quality than compared to lumber with 25%.

In contrast, Genichi Taguchi developed during the quality revolution in the 1970s quality loss functions which quantify the financial loss based on variation within the product quality [12]. Thus, if your product attribute is deviating from the target you will experience loss (Figure 10, right). And this loss is experienced after the shipment of the product in form of future need of repair or replenishment of the product and even damaged reputation of your company.

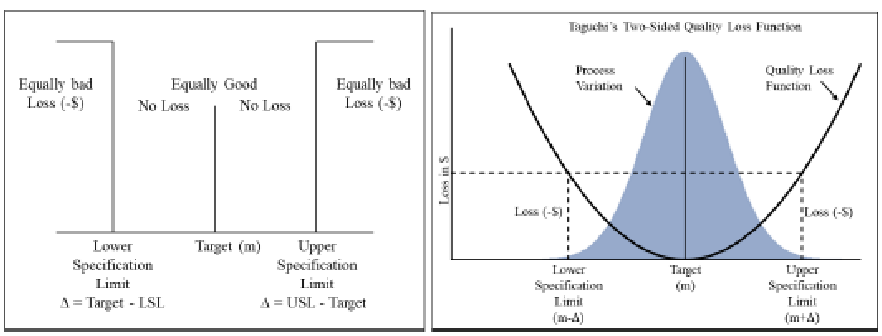


***Figure 10.** Traditional view of loss (left) and Taguchi's view on loss (right).*

Taguchi specified three different quality loss functions. Each quality loss function is designed for specific process variables or product attributes (Table 3). Furthermore, you can either compute the loss based on each value individually or as an average loss per unit for the whole data set. However, try to build an optimal loss function for your specific case.

Table 3. Overview Taguchi Quality Loss Function [12].

Type	When to use?	Loss Function	Cost Coefficient k
Nominal-the-best (Figure 11)	Variables with two specification limits	Individual Loss: $L = k \times (y - m)^2$ Average Loss per Unit: $L = k \times (\sigma^2 + (\bar{y} - m)^2)$	$k = \frac{A_0}{\Delta_0^2}$
Smaller-the-better (Figure 12, left)	Variables which should be small as possible, theoretically zero.	Individual Loss: $L = k \times y^2$ Average Loss per Unit: $L = k \times (\sigma^2 + \bar{y}^2)$	$k = \frac{A_0}{USL^2}$
Larger-the-better (Figure 12, right)	Variables which should be large as possible	Individual Loss: $L = \frac{k}{y^2}$ Average Loss per Unit: $L = k \times \frac{1}{n}\left(\frac{1}{y_1^2} + \frac{1}{y_2^2} + \cdots + \frac{1}{y_n^2}\right)$	$k = A_0 \times LSL^2$

Where:
L = Loss in Dollars
k = Cost coefficient \$ / (variable unit of measure) ^{2}
y = Individual value for your variable / attribute
$\bar{y}$ = Average of all values of the data set
σ^2 = Variance of the data set
m = Target of your process
Δ_0^2= Tolerance (e.g., USL – Target; Target – LSL)
A_0 = Consumer Loss
n = number of values in the data set

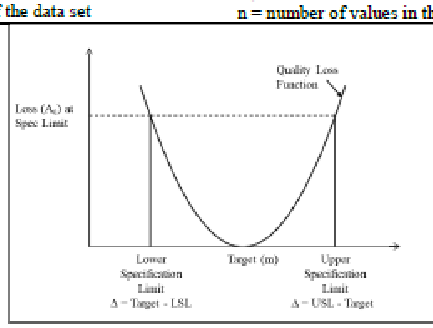


Figure 11. Quality Loss Function - Nominal-the-best.

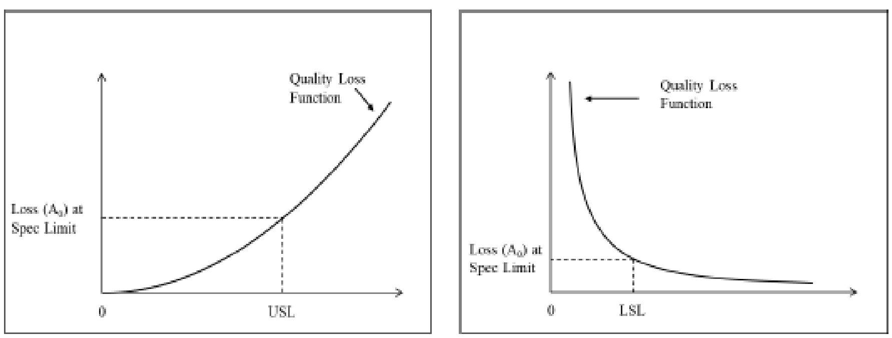


Figure 12. Quality Loss Functions - Smaller-the-better (left) and Larger-the-better (right).

Example for Taguchis' Quality Loss Function

Recall the example of a treated lumber production (target(m) = 25%, Tolerance (Δ) ± 3%). Assume you receive three batches of dried lumber from three different lumbermills for your wood treatment production. Now, ask yourself which lumbermill provides the best batch in terms of moisture content variation and thus financial loss. Since you have two specification limits Taguchi's nominal-the-best equation is used. Assume a customer loss at the spec limit of $18. Then cost coefficient k = 2 $/%².

Lumbermill A		*Lumbermill B*		*Lumbermill C*	
Moisture Content [%]	*Individual Loss [$]*	*Moisture Content [%]*	*Individual Loss [$]*	*Moisture Content [%]*	*Individual Loss [$]*
27.7	14.58	25.1	0.02	26.2	2.88
22.3	14.58	23	8	24.7	0.18
25.5	0.5	27.1	8.82	22.2	15.68
25.3	0.18	27.1	8.82	23.1	7.22
24.2	1.28	22.1	16.82	24	2
27	8	24.3	0.98	25.1	0.02
27.5	12.5	24.5	0.5	26.8	6.48
27	8	23.2	6.48	25.6	0.72
28.8	28.88	23.9	2.42	25.2	0.08
25.4	0.32	24.5	0.5	23	8
Total Loss	*$ 88.82*	*Total Loss*	*$ 53.36*	*Total Loss*	*$ 43.26*

Based on the results using the individual loss equation lumbermill C would provide your company with the best batch of lumber.
Nevertheless, to really see how variation influences the loss you could compare all three lumbermills based on the computed average loss per unit, which respects the variation of the data set.

Lumbermill A		*Lumbermill B*		*Lumbermill C*	
Average	*26.07%*	*Average*	*24.48%*	*Average*	*24.59%*
Variance	*3.66%²*	*Variance*	*2.66%²*	*Variance*	*2.22%²*
Average Loss per Unit	*9.61 $/Unit*	*Average Loss per Unit*	*5.87 $ / Unit*	*Average Loss per Unit*	*4.77 $ / Unit*

Thus, the lowest average loss per unit is experienced with lumber from lumbermill C.

Map the Sources of Your Variables Variation

After visualizing your variation of the 'critical few' process variables you must identify all possible root-causes. For this, apply Kaoru Ishikawa's philosophy or organized brainstorming via the so-called Fishbone chart or Ishikawa diagram (Figures 13 and 14) to categorize all sources of variation for the key product attributes and your process variables. Categorize usually are material, method, machine, man, and environment but of course can vary from case to case.

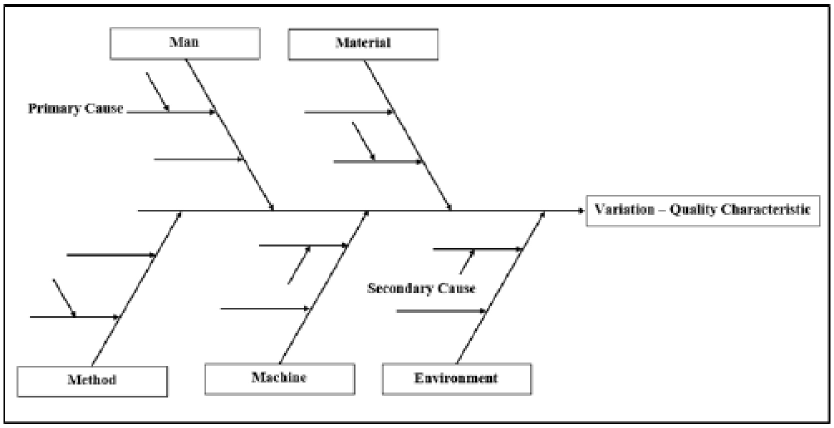


Figure 13. Typical Ishikawa diagram, a visualization and knowledge organization tool.

How to construct an Ishikawa-diagram:

- Form a team with people who have distinct knowledge about the machinery or process, e.g. machine operators or process engineers
- Place the main problem in the box on the right
- Generate and clarify all potential sources of variation via brainstorming as a team
- Categorize all sources / process variables into related groups
- Give those groups names
- Place the process variables on the appropriate bones of the Ishikawa diagram
- Combine each bone in turn, insuring that the process variables are specific, measurable and controllable. If they are not, "explode" the process variable until the ends of the branches are specific, measure, and controllable.

Tips:

- Identify causes and not symptoms.
- Publish diagrams in your company to get external perspective on the problem.

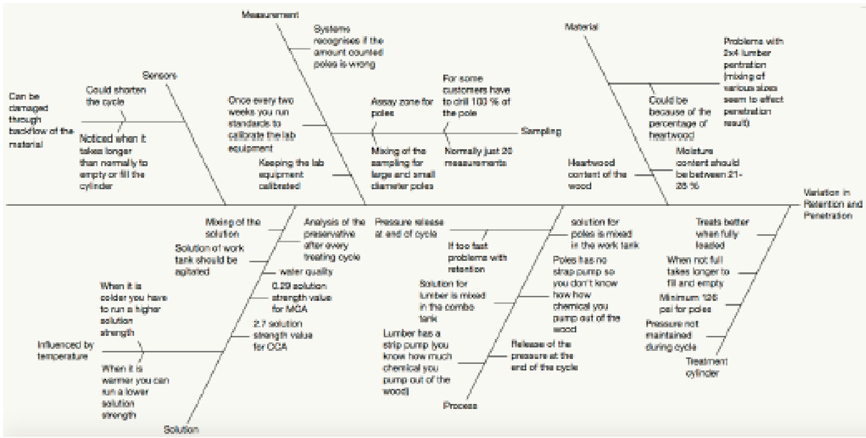


Figure 14. Ishikawa diagram for "Variation of Retention and Penetration of Treated Wood" [13].

Prioritize the Sources of your Variables' Variation

Pareto charts are visual aids for prioritizing the key process variables contributing to the main problem. Adapted from the "80/20-rule" invented by the Italian economist Vilfredo Pareto 80% of the variation in a process origin from 20% of the causes [14]. The identified causes for the problem are represented by bars on the horizontal axis; the cumulative contribution by the causes are represented on the vertical axis via a line (Figure 15) Pareto charts enable improved decision making for managers and engineers regarding tackling major causes for solving the main problem. Develop schemes which allows the worker to document the identified source or cause for variation. Later plot your knowledge / data to see which problem arises more frequently and start improving the most prominent problem.

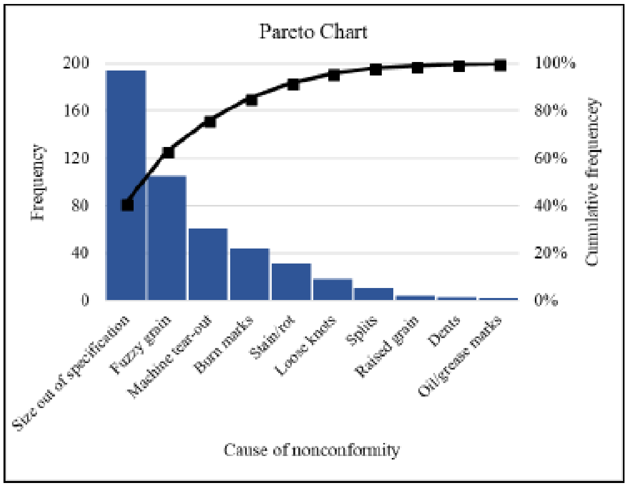


Figure 15. Pareto chart for causes of nonconformity of a wood product [15].

Continuous Improvement – The Plan-Do-Check-Act-Cycle

Most important for your Companies' success is to see continuous improvement as a long journey not a quick sprint. All entities of the company, starting with the upper management and ending with the machine operator, must pull in the same direction, be open for improvement opportunities and most importantly stay persistent. Form an inclusive, healthy, and constructive management culture within the company to create the foundation necessary for the successful implementation of continuous improvement. Use the philosophy of the *Plan-Do-Check-Act-Cycle* for every main cause for variation in your process variables (recall Pareto chart) to ensure your path on continuous improvement (Figure 16) [16].

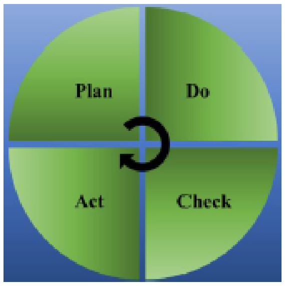


Figure 16. Plan-Do-Check-Act-Cycle.

Plan-Phase – Assess the current state of your process, i.e. current issues and deficiencies and develop a plan with countermeasures or actions which will improve the situation. Think and document the potential output of your countermeasures and changes. Make rather small changes to detect their impact more easily and make them more predictable / measurable.

Do-Phase –Execute your plan through implementing the countermeasurements in your process; e.g. *changed line speed, resin mix, drying temperature, or chamber pressure*. Document measurable statistics to visualize the impact of your counter actions.

Check-Phase – Evaluate the data results gathered from the *Do-Phase* and compare the results with your predicted outcome from the *Plan-Phase*. Use charts (e.g. control charts) to compare the results of multiple PDCA-Cycle runs.

Act-Phase – Act based on the results you gathered from your evaluation in *Check-Phase*. For positive results, i.e. improvements of the current standard setting of your process, adopt the countermeasures in your standard. For negative results, no improvements were made through the changes, keep the current standard without changes. Now, if you think improvements can still be made with your process start the cycle with the *Plan-Phase* again. Avoid making unplanned adjustments in the *Act-Phase*.

Getting rid of Bottlenecks – Theory of Constraints

Eliyahu M. Goldratt developed Theory of Constraints (TOC) to provide a thinking concept on how to tackle material or managerial limitations in manufacturing to greatly improve the systems performance [4]. These production limitations, bottlenecks, essentially constrain the process execution and as a result restrain the overall success of the enterprise [17]. A perfect enterprise would have no constraint and would make infinite profit [17]. Therefore, in TOC the success of an organization is based on how well all processes work together. This theory provides a five-step approach to solve the constraints individually and implements an additional way for continuous improvement of a system [4, 18, 19]

At first, the manager or engineer should start with (1) *identifying the system's constraint(s)*. The choice of constraint should be based on the constraints impact on the performance of the production. Constraints can be either physical, for example limited machine capacity or material variation or based on policy. Policy constraints can either be created from poor process methodology or by flawed design of regulations and rules in an organization. After the constraints identification there should be a discussion on (2) *how to exploit the system's constraint(s)*. Physical constraints should be used as effectively as possible. In contrast, a flawed policy should be eliminated and replaced with an improved new policy. (3) *Subordinate everything else to the above decision,* for achieving maximum success with the current production environment. By subordinating all resources to the main constraints need allows to maximize its output and essentially improve the total systems performance. This is possible since non-constraint resources have productive and non-productive capacities; optimal used non-constraint resources have no impact on the performance. If the identified (1) and exploited (2) (3) constraints are still existent it is crucial to (4) *elevate the system's constraint(s)* to generate more company profit. Elevating means to find actions to improve the systems overall performance. For example, if resource (machine) capacity is limiting the production output buying another machine to gain increased production capacity. Thus, another constraint in the production will arise and will form the new constraint - (5) *if a constraint was broken in a previous step, go back to step 1*. Step 5 implies that TOC should be seen and executed as a continuous improvement process; inertia should not allow to restrict the performance of the enterprise.

Conclusion

Start your journey of continuous improvement by seeing variation as a chance. A chance to improve your current performance. Design a secure data system for storing all you sensorics data. Start by mapping your production, identify your 'critical few' process variables and product attributes. Start your analysis by describing your data; plot a histogram how does the distribution look like? Use statistical methodologies to visualize (e.g., control charts) and to quantify (e.g. Taguchi's quality loss function) the variation of your process variables. Identify all potential sources which cause variation and prioritize the most impactful. If you do this, you are on a good way to improve your company. And finally, continuous improvement is a journey which will never end, stay persistent.

You can find more information either here www.spc4lean.com or the cited literature.

Literature cited

1. Singh, J. and H. Singh, *Continuous improvement philosophy – literature review and directions*. Benchmarking: An International Journal, 2015. **22**(1): p. 75-119.
2. Bhuiyan, N. and A. Baghel, *An overview of continuous improvement: from the past to the present*. Management Decision, 2005. **43**(5): p. 761-771.
3. Young, T.M., *Process optimization by means of statistics*, M. Weigl, Editor. 2015, Holzforschung Austria: Vienna, Austria.
4. Srinivasan, M.M., *Building lean spply chains with the theory of constraints*. 2011: McGraw-Hill Education.
5. Institute, M.R., *Emission Factor Documentation for AP-42 Section 10.8 - Wood Preserving - Final Report*. 1999.
6. Lebow, S.T., *Wood Preservation*, in *Wood Handbook - Wood as an engineering material*. 2010, Forest Products Laboratory: Madison, Wisconsin.
7. Wheeler, D.J. and D.S. Chambers, *Understanding statistical process control* 2ed. 1992, Knoxville, TN: SPC Press.
8. Young, T.M., B.H. Bond, and J. Wiedenbeck, *Implementation of a real-time statistical process control system in hardwood sawmills*. Forest Products Journal, 2007. **57**(9): p. 54-62.
9. Western Electric Company, *Statistical quality control handbook*. 1 ed. 1956, Indianapolis, Indiana: Western Electric Co.
10. Nelson, L.S., *Technical aids*. Journal of Quality Technology, 1984. **16**(4): p. 238-239.
11. Young, T.M. and P.M. Winistofer, *SPC: Statistical Process Control and the forest products industry*. Forest Products Journal, 1999. **49**(3): p. 10.
12. Taguchi, G., S. Chowdhury, and Y. Wu, *Taguchi's quality engineering handbook*. 2005, Hoboken, New Jersey: John Wiley & Sons, Inc.
13. Hamernik, J., *Sources of variation of treated wood*. 2018, Fachhochschule Salzburg University of Applied Sciences.
14. Wilkinson, L., *Revising the Pareto Chart*. The American Statistician, 2006. **60**(4): p. 332-334.
15. Leavengood, S. and J. Reeb, *Part 3: Pareto analysis and check sheets*, in *Statistical process control*. 2002, Extension Service, Oregon State University: Corvallis, Or.
16. Wilson, L., *How to implement lean manufacturing*. 2010: McGraw-Hill.
17. Blackstone, J.H., *Theory of constraints - A status report*. International Journal of Production Research, 2001. **39**(6): p. 1053-1080.
18. Goldratt, E.M., *What is this thing called theory of constraints and how should it be implemented?* 1990: North River Press.
19. Rand, G.K., *Critical chain: The theory of constraints applied to project management*. International Journal of Project Management, 2000. **18**: p. 173-177.

Appendix

Table 4. Factors for Control Charts [7].

n	A_2	D_3	D_4
2	1.880	-	3.268
3	1.023	-	2.574
4	0.729	-	2.282
5	0.577	-	2.114
6	0.483	-	2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
0	0.337	0.184	1.816
10	0.308	0.223	1.777
11	0.285	0.256	1.744
12	0.266	0.283	1.717
13	0.249	0.307	1.693
14	0.235	0.328	1.672
15	0.223	0.347	1.653
> 15	$\frac{3}{d_2\sqrt{n}}$	$1-\frac{3\,d_3}{d_2}$	$1+\frac{3\,d_3}{d_2}$

Table 5. Bias Correction Factors [7].

n	d_2	d_3	n	d_2	d_3
2	1.128	0.8525	19	3.689	0.7335
3	1.693	0.8884	20	3.735	0.7287
4	2.059	0.8798	21	3.778	0.7272
5	2.326	0.8641	22	3.819	0.7199
6	2.534	0.8480	23	3.858	0.7159
7	2.704	0.8332	24	3.895	0.7121
8	2.847	0.8198	25	3.931	0.7084
9	2.970	0.8078	30	4.086	0.6927
10	3.078	0.7971	35	4.213	0.6799
11	3.173	0.7873	40	4.322	0.6692
12	3.258	0.7785	45	4.415	0.6601
13	3.336	0.7704	50	4.498	0.6521
14	3.407	0.7630	60	4.639	0.6389
15	3.472	0.7562	70	4.755	0.6283
16	3.532	0.7499	80	4.854	0.6194
17	3.588	0.7441	90	4.939	0.6118
18	3.640	0.7386	100	5.015	0.6052

VITA

Christoph Metzner was born in Traunstein, Germany, and he is currently a Master student at the University of Tennessee, Knoxville. He attended the University of Applied Sciences in Rosenheim, Germany, where he earned his Bachelor of Engineering in Wood Technology. He conducted "Investigations about fall protection of mechanical t-joints in plastic windows" for his bachelor thesis. The results of his thesis were later used as the foundation for a novel test guideline and an industrial norm.

Christoph is currently pursuing a master's Degree in Forestry with a concentration in Bio-Based Products and Wood Science and Wood Technology in the Department of Forestry, Wildlife and Fisheries under the advisory of Dr. Timothy M. Young. Based on a joint venture program between the University of Tennessee and the Salzburg University of Applied Sciences Christoph is pursuing a second Master of Science from SUAS.

He plans to graduate from the University of Tennessee with a "Master of Science" degree in Forestry (Bio-materials concentration) in December 2018 and a second Master of Science degree ("Diplom-Ingenieur") from SUAS in September 2018.