APPLICATION OF PRINCIPAL COMPONENT ANALYSIS AND ARTIFICIAL NEURAL NETWORKS IN THE DETERMINATION OF FILLER DISPERSION DURING POLYMER EXTRUSION PROCESSES

Jian Yan

A Thesis in the

Department of Mechanical and Industrial Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Applied Science (Mechanical and Industrial Engineering) at

Concordia University

Montreal, Quebec, Canada

August 2003

Your file Votre référence
ISBN: 0-612-83888-9
Our file Notre référence
ISBN: 0-612-83888-9

ABSTRACT

Application of Principal Component Analysis and Artificial Neural Networks in the Determination of Filler Dispersion during Polymer Extrusion Processes

Jian Yan

Mineral filler-reinforced polymer is an important family of polymers designed to achieve high mechanical impact strength. The state of mineral filler dispersion in a polymer matrix strongly affects the mechanical properties of the product and is an important information for the extrusion-based fabrication process. In this work, a measurement system consists of two ultrasonic sensors, three pressure sensors, a thermocouple, and an amperometer of the extruder motor drive were used to monitor the extrusion of a calcium carbonate powder-filled polypropylene system. Three principal components, most correlated to the state of filler dispersion, were extracted from the data set collected by the multiple sensors and fed as inputs to an artificial neural network model designed to determine the dispersion state of the filler. By using this approach, one is able to achieve an accuracy of better than 0.05 on the dispersion index. This work has demonstrated the feasibility of combining our multi-sensor monitoring system with prinicipal component analysis and artificial neural networks for on-line determination of mineral-filled dispersion in polymers.

ACKNOWLEDGMENTS

My foremost thanks go to my co-supervisors Professor Mingyuan Chen and Dr. Cheng-Kuei Jen. I thank them for their guidance and supports throughout my graduate studies. They advised and helped me in various aspects of my research. Also, their energetic and efficient working style has influenced me greatly to perform as a professional. I appreciate the financial support from the Concordia University Graduate Fellowship in my M. A. Sc. program.

I am grateful to my co-supervisor Dr. Zhigang Sun, who introduced and brought me into the field of ultrasound. His visionary thoughts, consistent encouragement and critical comments were crucial for my completion of this project. Without his assistance, this thesis would not have been possible. Besides, I sincerely thank Industrial Material Institute/National Research Council Canada (IMI/NRC) for providing the advanced instruments and equipments that allowed me to effectively proceed with my research.

In addition, I would like to acknowledge Dr. Jacques Tatibouët for his valuable assistance in polymer extrusion processing. My appreciation is extended to other members in Ultrasonic group of IMI/NRC, Dr. Yuu Ono, Jean-François Moisan, Harold Hébert, Christian Corbeil, Yuanbei Zhang, Donghai Zhao and Makiko Kobayashi for their kindly help and comments. Also I enjoyed all the interesting discussions we had on various topics and had lots of fun being a member of this active group.

Next, I wish to thank with all my heart to my wife, Ke Zhong, who has provided emotional, spiritual and financial support throughout the last two years. Her love and sacrifice encouraged me to focus on my graduate study. Last but not least, I thank my parents and my sister for always being there when I needed them most, and for supporting me through all these years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1 Background

Polymers are widely used in today's industry due to wide range of properties exhibited and ease of processing. A category of polymers that draws great interests is (mineral-) reinforced polymer, including calcium carbonate ($CaCO_3$) filled and fiber-filled polymers and more recently, polymer nano-composite materials. Basically, these fillers or additives are used to alter and tailor, at low cost, the properties of the matrix polymer, such as impact strength, flexural modulus, modulus of elasticity, and foamability, or are used to fulfill certain special functions, such as flammability resistance [4,13,47]. Generally, properties of filled or reinforced polymers depend on the size, shape and surface properties of the fillers [46]. For a mineral-filled polymer system, however, one of the most important parameters affecting mechanical performance is the state of dispersion that has been shown to correlate with the dynamic and mechanical properties of the composites [8,38,39,67,75]. The state of dispersion generally refers to the amount of particles in the composite in the form of primary particles (i.e. non-agglomerated particles) or to the amount of agglomerates within a given upper size limit [67]. Since material properties are often strongly affected by processes and also the improved understanding of processes will produce enhanced quality and increase profitability [58], one of the most challenging problems faced is how to obtain the timely information about the process in terms of the state of dispersion of minerals or particles.

Dispersion of agglomerates or particles has been applied in the polymer processing industry for over 50 years. Classical direct methods for obtaining and

evaluating dispersion are based on optical or scanning electron microscopy (SEM) observations of extruded specimens. Reported techniques also include surface roughness measurements [45], and electrical conductivity and dielectric measurements [3]. For thermoplastic systems, SEM of microtomed surfaces and image analysis methods [17,55] are commonly used to estimate surface based dispersion indices [22,67] and volume mean diameter dispersion indices [19]. However, both methods are time consuming and cannot be easily implemented on-line to meet the need of controlling quality of manufactured products.

Polymer extrusion is one of the most used industrial mass production processes of plastic products. In the process, polymer pellets are fed into the feed hopper, transported, mixed and blended by screws in the heated barrel section under elevated temperatures (about 150°C ~ 220°C), and then converted into melt at the extruder exit die that shapes the product. During this process, the complex thermo-mechanical history of the materials greatly affects the qualities of the final product.

Traditionally, the control of an extrusion process is achieved by measurements of melt pressures and barrel temperatures at different locations along the extruder for different screw rotation speeds and feed rates. However, these measurements cannot provide timely information on the properties of the materials except the running state of the machine. On most existing extruders, the extrusion process is controlled based on the information gathered on the status of different parts of the machine, for example, the barrel temperature, screw position and speed, etc. Conventional pressure and position sensors, thermocouples and amperometer are used. It is a machine-control strategy by which only the machine itself is directly controlled and the instantaneous knowledge of

the state of the melt is not taken into account by the control system. The consequence is that when some parts of the machine go wrong (e.g., screw misalignment or wear, variation of feed speed), or when material composition changes accidentally (e.g., polymer degradation caused by humidity in air, or batch-to-batch variation of composition), the machine will continue to operate in a programmed way without being able to cope with those changes that are not sensed by the above mentioned conventional sensors. This could result in a large number of defective products before the problem gets noticed and corrected, causing significant production and material losses.

Furthermore, these measurements cannot be carried out entirely without any problems. For example, the temperature of the channel wall gives no reliable indication of the effective temperature of the flowing melt, as there is always a radial temperature profile over the cross section of a flow channel. Since thermocouples can only be mounted on the barrel wall, the collected melt temperatures actually reflect the surface temperature of the material close to the barrel rather than the temperature profiles of the polymer melt along its thickness direction. Also, pressure is no accurate indicator; since a molten plastic zone at the heated wall may cause a significant reduction of melt pressure, whereas more solid material is located in the core. Besides, quality of products is often evaluated at laboratory off-line using samples collected from the extrusion lines. The time lag between collecting the samples and obtaining the final result of quality inspection may range from a few hours to several weeks, depending on the complex extent of the blend and lab availability. In some cases, long delays are encountered before the analysis results are obtained. In some cases, serious raw material waste may occur due to the delay of laboratory analysis results.

Ultrasound is sensitive to both processing conditions and material mechanical properties and can be used to pinpoint, in real-time, the defects occurred during either the processing or material conditions. Combining ultrasonic sensing technology with existing thermocouples, pressure sensors and amperometers will provide an extra edge in assuring consistent good quality parts. It is expect that ultrasonic monitoring can provide useful information for the design and optimization of polymer extrusion processes.

Ultrasonic techniques have a number of attractive features for measuring melt properties. For example, they have fast response, are non-destructive, and can be implemented in a non-invasive manner. Also, since ultrasound can penetrate through the entire flow channel of polymer melt, it can measure the properties at a relative large volume of polymer melt, in contrast to some existing optical (infrared, fluorescence, visible light) methods which usually measure at a surface layer only. These advantages were already recognized more than 30 years ago, when the first suggestions of using the technique for monitoring polymer-processing operations were made [29]. After then, attention turned to multiphase systems, more specifically, to attempts to relate ultrasonic attenuation with degree of dispersion of fillers in polyethylene (PE) and polypropylene (PP) melts [9,10,18]. Also, much work is specifically aimed at various industrial applications and has demonstrated that on-line measurements of sound velocity and attenuation are powerful tools for process monitoring [11,51,52,65,69,76]. Recently, a number of papers have been published on ultrasonic monitoring of fundamental extrusion performance, such as progress of melting and mixing [68], barrel and screw wear [34], and single-screw compounding [71]. The latest development of high temperature ultrasonic buffer rod probes [20,33,70] has been a key contribution to the recent advances

in ultrasonic monitoring of polymer processing. As far as monitoring of dispersion is concerned, ultrasonic techniques, developed recently, have demonstrated the capability of monitoring in real time the dispersion state of calcium carbonate in polypropylene in an industrial environment [72].

At the same time, ultrasonic techniques have found numerous applications in characterization of polymers, both in solid and molten states [61,62,74]. Ultrasonic velocity and attenuation were measured for various commercial polymers and it was suggested that ultrasound measurements could be used for quality control or for purposes of material characterization [28], in measurement of elastic properties of polymers [49,50] or in study of the crystallization process [24,73]. All these applications indicate that ultrasonic techniques open a way toward real time quality monitoring and possibly of closed-loop control of compounding processes.

In this work, ultrasonic velocity and attenuation of polymer melt together with traditional temperature and pressure measurements were collected and investigated at an instrumented die during a twin-screw extrusion process as part of the data of the $CaCO_3$-filled polypropylene extrusion process. Ultrasound signals, more specifically ultrasonic attenuation (α_{us}) and velocity (V_{us}), are affected by filler dispersion state (D_x), filler concentration (C_f), the type of filler (F_{type}), melt temperature (T_{melt}), pressure at ultrasonic probe location (P_{us}), and flow state. The melt temperature, pressure, and flow state at the probe location are affected by the temperature profile (T_{prfl}), the barrel temperature along the extruder (i.e., temperatures at each barrel section of extruder, T_{barrel}), extruder screw configuration, screw rotation speed (V_{RPM}), material feeding rate (Q_{feed}), and the properties of the matrix material polypropylene (PP) as well. Q_{feed}, T_{prof}, screw rotation

speed (V_{RPM}), F_{type} and C_f can be controlled by practitioners (called controllable variables here) and have strong influence on dispersion index. These studies revealed that some of these measured and controlled parameters are correlated. Also, due to tedious and time-consuming SEM sample analysis, a very limited number of analyzed samples were used. The challenge one had to face was how to use the available information generated by this complex multivariate system to effectively establish a relationship between the dispersion index and the measured and controlled variables for the on-line monitoring purpose.

Due to the lack of understanding of the interaction between ultrasound and filled polymers in flowing state, so far, there are no established theoretical relations available between the measured variables and the state of filler dispersion. Artificial neural network (ANN) approach provides a way for modeling this relationship without the need of thorough understanding of the process itself and therefore was chosen in this study. On the other hand, not all variables are necessarily equally important or relevant to the determination of dispersion state. There was a demand of reducing the information redundancy and extracting most informative variables. The application of principal component analysis (PCA) allowed finding several uncorrelated feature parameters that are most sensitive to the filler dispersion state. There has been work on the use of neural networks and ultrasound for quantifying the dispersion of mineral filler in a polymer [7]. This thesis examines a way that combines PCA and ANN to identify the correct correlation between the on-line multivariable measurement data and the dispersion index and builds a model that can be applied on-line to predict the dispersion index in real time from the known measurements.

1.2 Objectives

The purpose of this study is to adapt and combine PCA and ANNs, to analyze and interpret multivariable measurement data, to extract the most important or informative factors relevant to the dispersion index measurement, and to build a model that predicts the dispersion index of polymer process.

1.3 Outline of the thesis

This thesis is organized as follows: Chapter 2 describes basic concepts of polymer extrusion processes and filler dispersion. Chapter 3 briefly presents the theoretical basis of ultrasonics and experimental setup of an ultrasonic data acquisition system. In Chapter 4, an overview of PCA is given and in Chapter 5 a principal knowledge of Neural Network is outlined. Chapter 6 focuses on the application of PCA and ANNs for estimation of filler dispersion. A conclusion of this work is presented in Chapter 7.

Chapter 2 Polymer Extrusion and Filler Dispersion

Extrusion is the most used, and perhaps the most important method of plastic fabrication today. The commercial process of polymer extrusion is the conversion of a raw material, usually in the form of pellets or powder, into a finished product or part by forcing melting it and through an opening of a die. The process consists of feeding a molten state polymer, under pressure, through a die, producing a continuous cross-section or profile. Mineral filler-reinforced polymer is an important family of polymers designed to achieve higher mechanical impact strength, superior to non-reinforced polymer. The state of mineral filler dispersion in a polymer matrix is important information for the fabrication process (i.e., polymer extrusion) because of its strong effect on the mechanical properties of the product. This chapter will introduce polymer extrusion process and mineral filler dispersion. Polymer extrusion processes have been extensively studied for decades and described extensively in the literature. In order to provide the readers of this thesis with some basic knowledge about the process, the following two sections outline some key aspects of polymer extrusion and experiments mostly excerpted from literature. For more detailed information, the reader is referred to [4,46,47,54].

2.1 Polymer extrusion equipment

The main sections of an extruder are shown in Fig. 2.1. The core turning screw can be imagined as trying to unscrew itself backward out of the barrel full of material. Since it cannot go back because a thrust bearing holds it in place, it pushes material and forces material towards the exit end. The die (not shown in this figure) at the exit end acts as a

resistance. The longer and smaller it is, the more the screw must work to push the material out (more horsepower required of the motor that turns the screw).

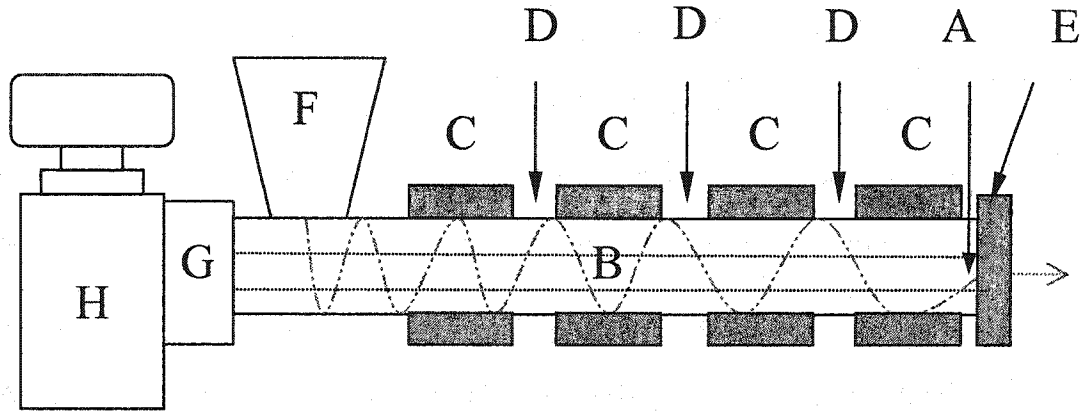

Fig. 2.1 Schematic diagram of polymer extrusion. A: Screws. B: Barrel. C: Heater band. D: Sensors well. E: Exit die. F: Hopper. G: Main thrust bearing. H: Gear box.

1) **Feed Hopper** - The hopper contains a large amount of feed material in the form of powder or pellets. This is gravity fed onto the upper surface of the exposed screw that continuously pushes material into the barrel between the flights.

2) **Barrel** - The barrel is normally heated to melt the polymer or initiate crosslinking. The barrel is of constant inner diameter and has heavy walls to withstand high pressures. A heating element is usually wrapped around the outside diameter of the barrel. The barrel runs the entire length of the screw from the hopper, where its upper side is fitted to the hopper, and to the die where it narrows, with a shaped opening through the die. The range of the inside diameter of the barrel is from 3/4 to 24 inches.

3) **Extruder Screw** - The screw is a helix that feeds polymer along the barrels. The shape of the screw or right-hand helix on a rod that turns at certain number of revolutions per minute (RPM) determines the speed at which the material feeds and the

pressure attained in the barrel. The screw is so named because its general shape is a helix as in a fastener. The continuous central rod of the screw is called the core. The diameter of the core is a major factor in determining the pressure in the barrel. The *L*/*D* ratio is the characteristic used to describe the size of the screw, where *L* is the total length of the screw and *D* is the inner diameter of the barrel. The shortest extruders have a ratio of 12, the longest 42. The conventional plasticating or single stage screw has three basic regions: the feed, transition and metering sections. The feed or solids conveying zone is one that transports the feed away from the hopper into the enclosed barrel. The feed is still in a solid powder or pelletized state and the screw has deep flights in this section. The transition zone is where compression occurs as the core diameter increases and the melting process takes place due to friction and heat from the barrel. The depth of the flights decreases along this section because the root changes size. The metering zone is the last section near the die so the polymer is molten and the depth of the flights are shallow and rather constant. There are several types of twin-screw extruders, such as intermeshing, or non-intermeshing, corotating, counterrotating or coaxial. They are characterized by more design variables than the single screw, but some can be used in very different applications. The details of different extruders are not discussed here.

4) **The Extruder Die** - The extruder die has an opening for the product to take shape. The die must withstand the high temperatures and pressures exerted on it by the polymer being forced through it. The polymer adopts the shape of the flow channel of the die. The pressure built up, called die-head pressure, depends upon the properties of the polymer, temperature of the polymer, the shape of the die and the flow rate through the

die. Most polymers experience some form of swelling upon exiting the die. This will be further discussed in section 2.2.

2.2 Polymer extrusion process

The largest volume of polymers is probably processed by means of extrusion. The extruder is the main device used to melt and pump materials through the die for shaping purposes. Basically, there are two types of extruders: single- and twin-screw. The single-screw extruder basically consists of a screw that rotates within a metallic barrel. The main function of the plasticating extruder is to melt solid polymer and to deliver a homogeneous melt to the die at the end of the extruder. The extruder can also be used as a mixing device, or a reactor. Also, there are as many twin-screw extruders as single-screw extruders in use today. Many different configurations are available, including corotating and counterrotating screws and intermeshing and nonintermeshing screws. Theses extruders are primarily adapted to handling hard-to-process materials and are used for compounding and mixing operations. The polymers used are typically thermoplastics. The pellets are compressed in the channel of the screw and then dragged forward by screw flights between the pellets and the barrel. Melt is caused by the heat generated by sliding friction at the barrel surface and transferred from the heated barrel. The opening in the die is the guide after which the extrudate takes its final form. There is a feed hopper for materials inlet at beginning of the cylinder and a die at the end. One of the real machines is shown a photo in Fig. 2.2.

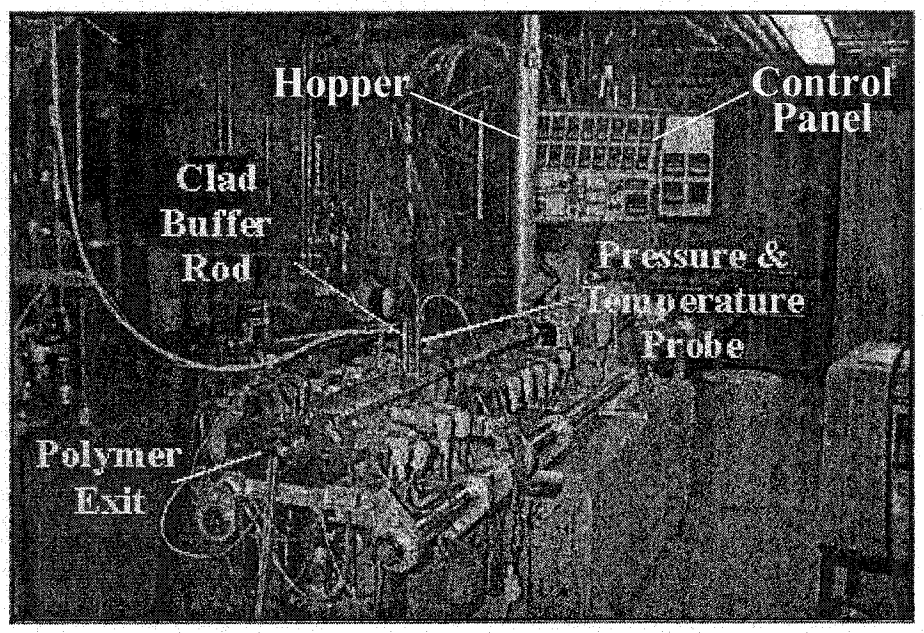

Fig. 2.2 Picture of a 34mm twin-screw extruder from Leistritz, Nurnberg, Germany

Most of the polymers used in extrusion are of high molecular weight and by nature are highly viscous when in the molten state. In order to process such materials the pump of the extruder must work under high pressures and temperatures. Thermoplastics are the predominant feed for extrusion processes. Due to the shearing action inherent in the screw feed mechanism, the process lends itself to dividing, heating up and melting extrudate. This does not exclude thermosetting polymers. Thermosetting polymers and elastomers can be fed, mixed with additives and crosslinking initiated by the heat in the barrel, usually completing the crosslinking after passing through the die. Examples of this are rubbers with vulcanizing agents and high-density polyethylene (HDPE) crosslinked by radiation.

When such a material is fed to the hopper, it is caught by the screws (see Fig. 2.3) and pushed through the barrel (see Fig. 2.4), where it gets hot and softens enough to continue and exit at the die. The heat is generated by friction as the screw turns in the

plastic mass. Thus, the main energy to melt comes from the motor as it turns the screw. More heat is often provided by external barrel heaters and/or preheated feed.

Fig. 2.3 Picture of twin screws

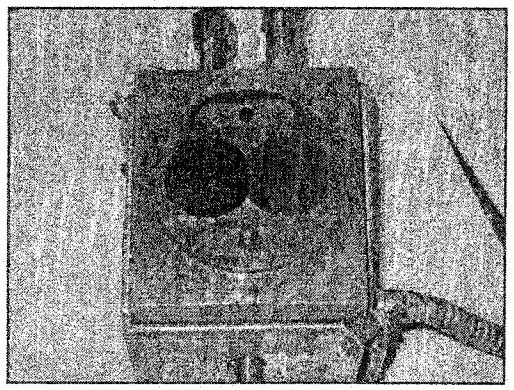

Fig. 2.4 Picture of a twin screw barrel with outside heating band

As the hot, soft plastic comes out of the die, it takes the shape of the hole it passes through – a long slit makes a film or sheet, a circular opening makes a rod, many small holes make filaments, etc. Once out of the die, the plastic must be cooled quickly (by air, water or contact with metal) and pulled away, rolled or cut up to the desired dimensions. For a continuous process of thin or very flexible material after cooling it is collected on rolls. For rigid material it is normally cut to lengths of up to 20 meters depending upon the transportation facilities available and collected in stacks.

Due to the combination of viscous and elastic nature of polymers there is some recovery by the polymer after passing through the die. The elastic nature of the polymer is to remember its shape, resist change and recover the shape it had prior to the distortion imposed by the die. The viscous nature of the polymer quickly accepts the shape forced upon it by the die without memory. The combination, viscoelastic behavior results in a swell after passing through the die. Since there is a considerable necking down from the

barrel size to the passage in the die, the polymer swells to partially return to its former shape. So the final product from the extrusion is not the size of the die passage but larger.

The final products of extrusion depend heavily upon the extrusion process. Products may be formed with solid cross sections or hollow cross sections. Those of solid cross sections such as angles, rods or strips are predominantly done in a uniformly controlled process. The barrel is maintained at a constant temperature to provide a constant melt. The screw speed is also constant so these result in a uniform flow rate. With conditions well controlled the die-swell will also be constant so that a product with consistent dimensions can be produced. For very exacting work or situations where normal controls are not adequate, sizing after the die is sometimes used.

Products with hollow cross sections such as tubes, cylinders and channels are also extruded. These require a die-core or mandrel that is the form around which the feed forms its shape. These are supported near the rear end of the screw by a spider of arms or by restricting the feed to enter a fraction of the circumference with the rest being filled by the support. Sizing is frequently used for hollow sections to fix the internal or external dimensions. The dimension that is not sized is determined by the polymer output rate, the haul-off rate and the size of the forming die used.

Forming flat sheet brings up a special set of problems. A small deflection in the die lips can cause large error in the thickness of the final sheet. Uniform heating of the die is difficult. Extrusion of a film is very similar to a sheet but the thickness variation due to deflection in the lips has even greater importance. Since the thinner films are more flexible, the unsupported gap between the die lips and the cooling must be reduced. To

form a tubular film by the blown method is different from the sheet method. The melted polymer is forced around an annular or crosshead die to form the film. Through the die air is pumped inflating the tube and providing some cooling as does an outer ring pumping cooling air over the surface of the expanded tubular bubble. A set of boards then guides the tube to a set of rollers (called nip rollers), the film is then wound flat. The tube is usually blown vertically avoiding the sagging effects of gravity.

2.3 Filler dispersion

Fillers are widely used in today's polymer industry because they can alter or enhance the properties of polymer in a flexible and easy-processing way. Materials used as particular fillers include wood flour (finely powdered sawdust), silica flour and sand, glass, clay, talc, limestone, and even some synthetic polymers. Because these inexpensive materials replace some volume of the more expensive polymer, the cost of the final product is also reduced. Among them, mineral fillers are often used to increase the mechanical properties.

Mechanical properties of small particle filled polymer composites are determined by various factors. These include the size, shape and fraction of fillers, interfacial properties and viscoelastic properties of the dispersed phase and the matrix phase, etc. In addition of these, state of dispersion of filler is one of the important factors to the mechanical properties, such as tensile and impact strengths. The impact strength increases as particle size decreases as long as agglomeration does not occur [46]. In other words, the mechanical properties are strongly related with the dispersion of mineral fillers during fabrication process. In general, impact properties are maintained if particle

with sizes in the micrometer or sub-micrometer range are used. The compounding step must however ensure that the particles can be wetted and distributed before they can agglomerate.

The state of dispersion usually refers to the amount of unagglomerated particulates that are present in a compound. To determine the state of dispersion, size and number of agglomerates are measured because these may act as defects in the composites due to their large size comparing with unit particles and aggregates, and relatively weak structure formed by particle-particle interaction forces. According to Suetsugu [67], dispersion index was defined as a function of area fraction of agglomerates

$$\text{Dispersion index: } D_x = 1 - \Phi_a \tag{2.1}$$

where Φ_a is the area fraction of agglomerates defined by

$$\Phi_a = \frac{\pi}{4A\Phi}\sum d_i^2 n_i \tag{2.2}$$

where A is the area of observation, Φ is the volume fraction of filler, d_i denotes diameter of agglomerate greater than a threshold value and n_i is the number of agglomerates with diameter d_i. Φ_a is a function of both the size and the number of agglomerates. Dispersion index varies between 0 and 1; 0 indicating worst dispersion where all filler particles remain in the form of agglomerates and 1, best dispersion where no agglomerate exists.

During the compounding operation, the dispersion is generally achieved in four steps: incorporation, wetting, agglomerate break-up and aggregates spatial distribution in the polymer matrix. Initially, for those ground minerals such as natural calcium

carbonate, talc and mica, there is no agglomerate structure as would be found in very fine carbon black. During the process of compounding, the ground particles are expected to be wetted and distributed into the polymer. At high concentration, compressive forces may be applied on the particles before the wetting and distribution are completed. If cohesive forces are sufficient, this can result in the formation of agglomerates. Compounding polymers with additives is generally obtained in corotating twin screw extruders. Fig. 2.5 shows the pictures of good and bad dispersion state respectively from the system of calcium carbonate filled into polypropylene matrix. These pictures were taken using optical microscopy. The black dots are calcium carbonate agglomerates.

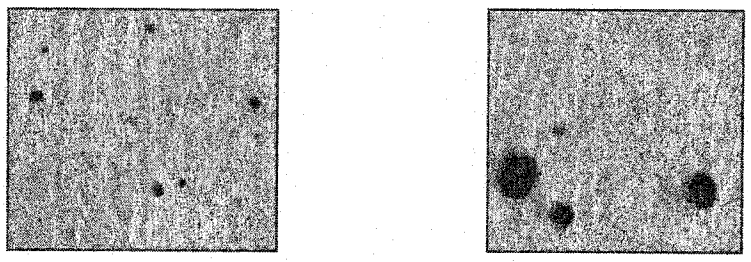

Fig. 2.5 $CaCO_3$ filled into polypropylene (microscopy)
Left: Good dispersion, agglomerate size less than 12μm
Right: Bad dispersion, agglomerate size around 150μm

2.4 Summary

Extrusion is one of the widely used industrial mass production processes to manufacture plastics products. It can be used at some stage in nearly all polymer processing operations. Basically, extrusion process consists of an extruder (single- or twin-screw), a die for shaping purposes and auxiliary equipment for cooling and profiling or pelletizing.

In the process, the polymer pellets are fed into the hopper, transported, mixed, and blended by the screws in a barrel. Heat generated by sliding friction at the barrel inner surface and transferred from the heated barrel causes the pellets to melt before the exit die where it takes the final shape. Depending on the extrusion process, the final products of extrusion can be solid cross section, hollow section, flat sheet, and tubular film. Also, the four key parts of an extruder, i.e. feeding hopper, barrel, screw and die are briefly described.

The properties and quality of materials incorporating fillers in their manufacture are affected by the state of dispersion of agglomerates and their mechanical properties, such as tensile and impact strength. It is known that the mechanical properties are affected by not only the size of agglomerates but also the number of agglomerates. The dispersion state refers to the amount of particles in the matrix that are in the form of agglomerates exceeding a given upper size limit. Here we use the relative area fraction, which defined the dispersion index as the ratio of the area of agglomerates exceeding certain size threshold over the detected surface area.

Chapter 3 Experiment Setup and Equipment

Several techniques, such as scanning electron microscope (SEM), optical microscopy, electrical conductivity and dielectric measurement, have been reported as used for the measurement of the dispersion state of certain filled polymer systems. Among them, SEM and optical microscopy are most commonly used in polymer industry. Both of them require tedious work in samples preparation and laboratory analysis, and are not suitable for in-line monitoring of extrusion process. In order to achieve timely control of product quality, a technique capable of providing real-time information on filler dispersion is highly desirable. Ultrasound has a long history of successes in industrial process monitoring and material characterization, owing to its robustness, fast response, non-destructiveness, non-invasiveness, and cost-effectiveness. Hence, ultrasonic technique was chosen in this study and it constitutes a part of an advanced multi-sensor monitoring system, in conjunction with other routine measurement instruments.

3.1 Setup of ultrasonic probes and measurement principle

Ultrasonics is a versatile technique that can be applied to a wide variety of material and process analysis applications. While ultrasonics is perhaps better known in its more common applications for thickness gauging, flaw detection, and acoustic imaging, high frequency sound waves can also be used to discriminate and quantify some basic mechanical, structural, or compositional properties of solids and liquids.

Ultrasonic technique is based on a simple principle of physics: the motion of any wave will be affected by the medium through which it travels. Thus, changes in one or more of four easily measurable parameters associated with the passage of a high frequency sound wave through a material - transit time, attenuation, scattering, and frequency content - can often be correlated with changes in physical properties such as hardness, modulus of elasticity, density, homogeneity, or grain structure. A significant advantage of ultrasonic measurement over other methods is that it can often be performed in process or on-line. Because sound waves penetrate through the test specimen, material properties are measured in bulk rather than just on the surface. In other words, ultrasonic technique can look inside of the material.

Ultrasonic measurements can be operated in reflection and transmission mode. In the reflection mode, the signal is transmitted and received by the same ultrasonic transducer (UT), while in the transmission mode, the signal is transmitted by one UT and received by another (see Fig. 3.1). The transmission mode has stronger signal and larger signal to noise ratio than reflection mode and was chosen in this application. When ultrasonic waves hit the boundary between two different media, some of the energy is transmitted though the boundary and the rest is reflected back. The received signal is then amplified and analyzed. A variety of commercial instrumentation is available for this purpose, utilizing both analog and digital signal processing. The setup of our multi-sensor monitoring system is presented in the next section. The ultrasonic velocity and attenuation coefficient can be calculated using the following equations.

$$Velocity: v = 2d / \Delta t \tag{3.1}$$

$$Attenuation: \alpha = (10/d)\log(A_2/A_1) \tag{3.2}$$

where d is the thickness of melt flow; Δt denotes the time period between two consecutive echoes; A_2 and A_1 represent the amplitude of two consecutive echoes (Figs. 3.1 and 3.2).

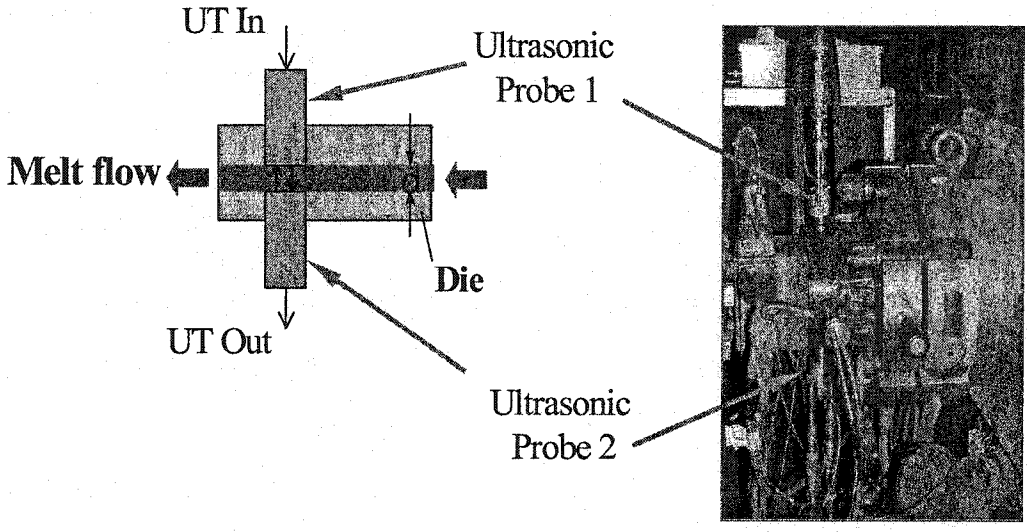

Fig. 3.1 Schematic diagram of ultrasonic measurement. Right is the picture of installation of two ultrasonic transducers at die when in measurement.

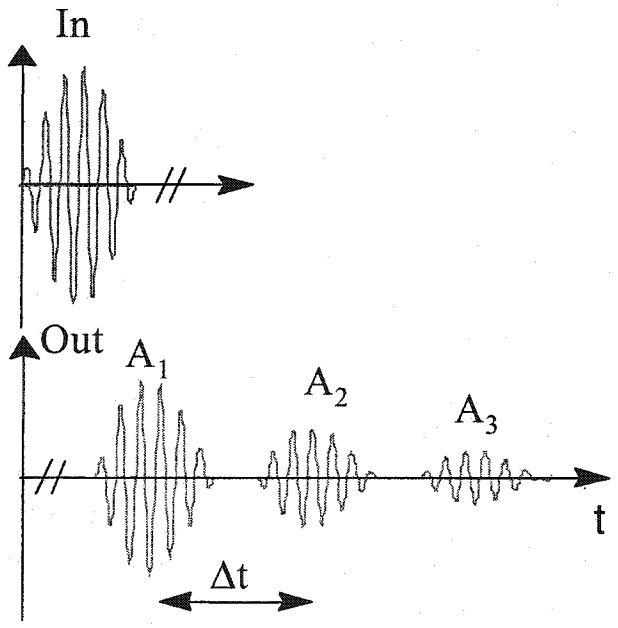

Fig. 3.2 Schematic diagram of ultrasonic signals sent and received by UTs.

3.2 Equipment

In this work, ultrasound, temperature, and pressure sensors as well as an amperometer of the extruder motor drive were used to monitor the extrusion of mineral-filled polymers under various experimental conditions in terms of filler type, filler concentration, feeding rate, screw rotation speed, and barrel temperature.

A Polypropylene (PP) 6631 from Himont Canada, with density $\rho = 890$ kg/m^3 and melt flow rate MFR = 2.0 dg/min, was used in this study. Two grades of calcium carbonate powders, Camel-Cal and Camel-Cal-ST, supplied by Genstar, with specific gravity of 2.71 were used as fillers. Both grades had a mean particle size of 0.7 μm, with 90% of the particles finer than 2 μm, and 100% finer than 7 μm. The grade with suffix ST was stearate-coated to allow for better particle dispersion than the non-coated one.

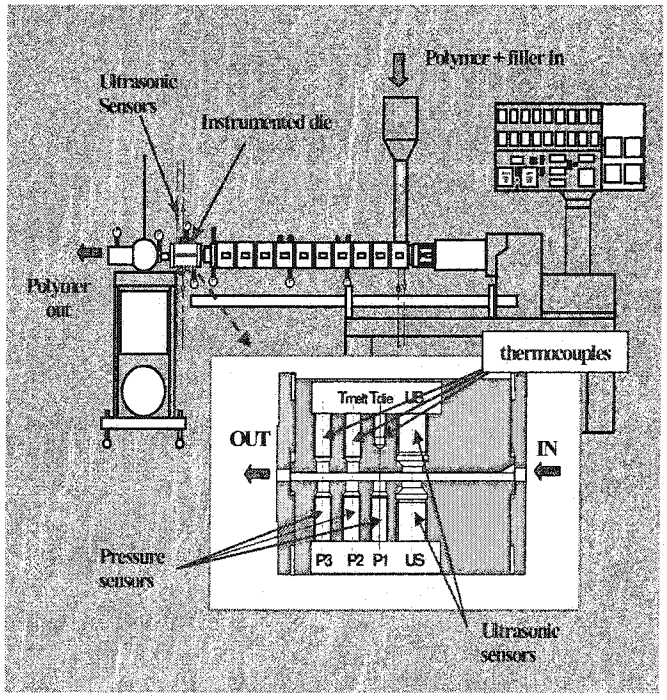

Fig. 3.3 Schematic layout of the extruder and instrumented die used in the experiments

A Leistritz 34-mm co-rotating intermeshing twin-screw extruder composed of 12 barrel sections with an L/D ratio of 42 was used. Polypropylene pellets and calcium carbonate powder were fed separately by two high precision gravimetric feeders at the same feed throat location (Fig. 3.3). A slit die instrumented with two 5-MHz ultrasonic sensors, a melt thermocouple, and three pressure transducers, all flush-mounted in the die slot, was used (Figs. 3.1 and 3.3). The die slot had a 3.0 mm high by 40.0 mm wide rectangular cross-section. The ultrasonic and pressure sensors were installed perpendicular to the wider surface of the slot. Two ultrasonic sensors were axially aligned on the opposite sides of the slot. During measurements, one of the ultrasonic sensors was used as a transmitter to send ultrasonic waves to the molten polymer. The other ultrasonic sensor was used as a receiver. The ultrasonic waves were reflected back and forth several times between the two ultrasonic sensors before completely dying out. As mentioned earlier, by measuring the time delay and amplitude difference between consecutive echoes, the ultrasound speed and attenuation coefficient in the molten polymer were calculated.

Chapter 4 Principal Components Analysis

One of the most important problems in data or signal processing is that a large amount of data is to be manipulated for information transmission, storage or processing. Among the information, not all are necessarily equally important or relevant to our research target. Also, redundant information exists in the measurement data due to some heavily correlated measurements. For instance, our study showed that the pressures P_1 and P_2 were highly correlated with a correlation coefficient of 0.98. For several decades, many statistical and mathematical methods have been developed in this field to efficiently extract useful information and established the discipline of chemometrics. Among them, principal components analysis (PCA) is a technique that specializes in reducing data size while maintaining as much information as possible. Basically it produces a set expression patterns known as principal components. Linear combination of these patterns can be assembled to represent the behavior of all samples in a given data set. It should be noted that PCA is not a widely used clustering technique. Rather, it is a fundamental tool of chemometrics to characterize the most abundant themes or building blocks that reoccur in many samples in the experiment. As a result, by using an algorithm that can identify the so-called principal components, one can leave out the rest of the data without significantly distorting the information born from the original data set. Beginning with the brief introduction of chemometrics, the following sections excerpted from related literatures focus on mathematical and application aspects of PCA. For more detailed information of PCA, the reader is referred to [31,36]

In this chapter, the following conventions will be used:

- Scalars will be represented in lower case, e.g. a, b, x,
- Vectors will be in bold lower case, e.g. $\boldsymbol{x}$, $\boldsymbol{t}$,
- Matrices will be bold upper case, e.g. $\mathbf{X}$, $\mathbf{S}$,
- The transpose of matrix $\mathbf{X}$ will be denoted by $\mathbf{X}'$,
- In 2-way data matrices, it is assumed that the rows of the matrix represent samples or observations and the columns represent variables, and
- A vector $\boldsymbol{x}$ is assumed to be a column vector; row vectors are represented by transposes, $\boldsymbol{x}'$.

4.1 Chemometrics

A general trend in analytical chemistry is to produce more and more data per sample. This is due to increasing analytical demands for improved specificity and sensitivity. It has been facilitated by developments in instrumentation and computer systems, making large amounts of data possible to produce and store with good economy. In order to make efficient use of the sophisticated analysis, it requires methods that can help analytical chemists to perform good experiments and to extract relevant information from the acquired data.

We first briefly introduce chemometrics. According to Wise et al. [78], chemometrics is the science of relating measurements made on a chemical system to the state of the system via applications of mathematical or statistical methods and the science of designing optimal experiments for investigating chemical systems.

Originally developed in analytical chemistry, many chemometric methods have been used in a wide range of chemical applications today. Some of them have been applied to problems in chemical engineering and chemical processes. The Chemometrics is data-based. The goal of many chemometric techniques is to estimate one or more properties of a system based on the production of a mode derived from the measurement data.

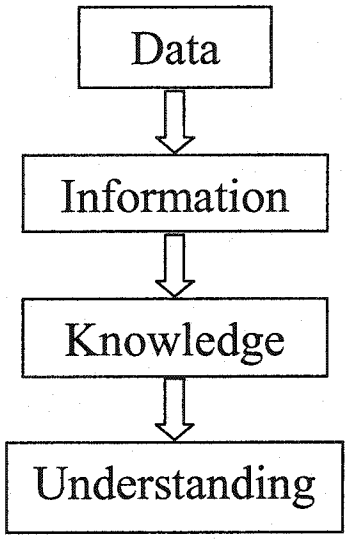

Fig. 4.1 The information hierarchy

Beyond simply obtaining a model that provides useful predictions, chemometric techniques can also be used to obtain insight about a chemical process. This is illustrated in Fig. 4.1. One typically starts with some data taken from a system in a tabular form consisting of measured variables on a number of samples or process measurements as a function of time. Without further processing, such a data set would be of little use. It is simply a record of measurements. If some calculations on the means and variances of the data are performed, and perhaps some trends or correlations between the variables can be obtained, then at this point some information on the system has been extracted. By

acquiring and studying additional data from the same system, one may obtain some knowledge on the system functions with fundamental principles from chemistry and physics to develop an understanding of the system. This understanding may lead to improvement in the system and in controlling the processes.

4.2 Background of PCA

Due to advanced instrumentation and frequent data recording, it is common for a process to have hundreds or even thousands of measured variable values. For instance, in a typical Fourier transform infrared spectroscopy (FTIR) spectrometer, the absorbance is measured at over 10,000 frequencies. Without adequate tools, it is impossible to interpret the data. In general, there are many correlated or redundant data in process measurements. They must be compressed in a manner to retain the essential information and are easy to display. Also, essential information does not come from an individual process variable but often derives from how the variables change with each other, i.e. how they co-vary. In this case the information must be extracted from the data. Furthermore, in the presence of significant noises, it is desirable to take signal averaging. Among the widely used multivariate statistical methods [2], principal component analysis (PCA) is an ideal tool for analyzing such data because of its ability to handle large numbers of highly correlated, noisy and redundant variables.

PCA is a tool in chemometrics for data compression and information extraction. Using this technique, a number of related variables are transformed to a set of uncorrelated variables [31]. PCA was originally developed in statistics by Pearson [48]. In geometry, PCA can be understood as the projection of a point swarm in an M-space to

a lower-dimensional subspace with P dimensions. PCA was further advanced by Hotelling [27] in the 1930s. Since then, applications of PCA have been found in many scientific fields, such as chemistry [56] and geology [16].

With the application of eigenvector decomposition to the covariance of the process variables, PCA finds combinations of variables or factors to describe major trends in a data set. They are useful in describing or predicting certain events or phenomena. And these combinations of variables are often more robust indicators due to the signal averaging aspects of PCA. Similar to other main multivariate techniques, PCA can be used to:

(1) provide insight into the process operation through preliminary analysis of the data;

(2) monitor process-operating performance through multivariate statistical process control methods;

(3) build predictive models for inferential control;

(4) identify dynamic models for large multivariable processes.

In this research, only the first function was applied to analyze the extrusion process data.

4.3 Some preliminaries

For an $m \times n$ data matrix **X**, the covariance between the data in the ith and jth column of **X** is calculated by:

$$s_{ij} = \frac{m\sum x_{ik}x_{jk} - \sum x_{ik}\sum x_{jk}}{m(m-1)} \tag{4.1}$$

where the summation is taken for k from 1 to m. Eq. (4.1) includes the mean-centered operation to the columns of **X**, i.e., each column is adjusted to have a zero mean by subtracting from every element the mean of the column. Hence, the covariance matrix **S** is given by

$$\mathbf{S} = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1n} \\ s_{12} & s_2^2 & \cdots & s_{2n} \\ \vdots & \vdots & & \vdots \\ s_{1n} & s_{2n} & \cdots & s_n^2 \end{bmatrix} \tag{4.2}$$

where s_i^2 is the variance of the ith variable, x_i , and s_{ij} is the covariance between the ith and jth variables. In matrix format, **S** can be given by:

$$\mathbf{S} = \frac{\mathbf{X}'\mathbf{X}}{m-1} \tag{4.3}$$

Note that the columns of **X** should have been mean-centered. Matrix **S** is symmetric and nonsingular. If a covariance is not zero, a correlation exists between the two corresponding variables. The strength of the correlation is represented by the correlation coefficient is: $\gamma_{ij} = s_{ij}/(s_i s_j)$.

In addition to the covariance matrix, correlation matrix is often utilized in PCA applications. In the correlation matrix, the column of the original data is "autoscaled". They are adjusted to zero mean and unit variance by dividing each column by its standard deviation. Eq. (4.3) can represent either a covariance matrix or a correlation matrix. Unless otherwise noted, it is assumed that the data is autoscaled prior to principal components analysis.

PCA uses some results from matrix algebra. A $p \times p$ symmetric, nonsingular matrix, such as the covariance matrix **S**, can be reduced to a diagonal matrix **L** by pre-multiplying and post-multiplying by an orthonormal matrix **U** such that:

$$\mathbf{U}'\mathbf{S}\mathbf{U}=\mathbf{L} \tag{4.4}$$

The diagonal elements of **L**, $l_1, l_2, \cdots, l_p$ are eigenvalues of **S**. The columns of **U**, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are *eigenvectors* of **S**, corresponding to the eigenvalues. $l_1, l_2, \cdots, l_p$ can be obtained from the solution of the following equation:

$$|\mathbf{S}-l\mathbf{I}|=0 \tag{4.5}$$

where **I** is the identity matrix. This equation produces a pth-order polynomial of l from which the values $l_1, l_2, \cdots, l_p$ are obtained. The eigenvectors may then be obtained by solving the following system of homogeneous linear equations

$$[\mathbf{S}-l\mathbf{I}]\mathbf{u}_i = 0 \tag{4.6}$$

Note that vector $\mathbf{u}_i$ should be normalized to length of 1. Therefore matrix **U** can be formed and is orthonormal, which means

$$\begin{cases} \mathbf{u}_i'\mathbf{u}_i = 1 \\ \mathbf{u}_i'\mathbf{u}_j = 0 \end{cases} \quad \text{for } i,j = 1, 2, \cdots, p. \tag{4.7}$$

Although the above mentioned direct solution of the characteristic equation is simple, it is difficult to solve directly for more than three variables due to the increasing difficulty of solving the equation (4.5). In this research, the Power method to find the

eigenvalues and eigenvectors was used. Detailed steps of the Power method can be found in any engineering numerical method textbooks such as [31].

4.4 Method of PCA

As mentioned in previous chapter, PCA is a tool for data compression and information extraction, which is especially useful for ill-conditioned data that has redundant information. In general, PCA includes the following steps [32]:

(1) Choice of the variables to be calculated in the analysis;

(2) Construction of the basic data matrix;

(3) Transformation of the basic data, mean-centering or autoscaling;

(4) Calculation of the covariance or correlation matrix;

(5) Calculation of eigenvalues and eigenvectors of the covariance or correlation matrix;

(6) Plotting or further analysis of principal components.

In this study, there were four pressure-related variables obtained from the monitoring system. Information redundancy exists in the data and one of the four pressure-related variables may be dominant. PCA is to find combinations of variables to describe major trends in the data.

From statistics point of view, the information of a process is carried by these major trends. For a given data matrix $\mathbf{X}$ with m rows and n columns, there are two ways to measure the overall variability of the data. They are:

1. The determinant of the covariance matrix, $|\mathbf{S}|$. This is called the generalized variance. The square root of this quantity is proportional to the area or the volume generated by the data.

2. The sum of the variances:

$$s_1^2 + s_2^2 + \cdots + s_n^2 = \mathrm{Tr}(\mathbf{S}) \text{ (The trace of } \mathbf{S}) \tag{4.9}$$

Either of them can be used in PCA. We utilized the second method in this research. PCA calculate the vectors that are a linear combination of the columns of **X** to describe the level of variability in the data. These vectors are the eigenvectors of the covariance matrix of **X**. From singular decomposition method, **X** can be written by a linear combination as:

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1' + \mathbf{t}_2\mathbf{p}_2' + \cdots + \mathbf{t}_k\mathbf{p}_k' + \mathbf{E}. \tag{4.10}$$

where k is less than or equal to the smaller dimension of **X**, i.e. $k \leq \min\{m, n\}$. **E** is the residual matrix. $\mathbf{t}_i$ are scores and $\mathbf{p}_i$ are eigenvectors of the covariance matrix **S**, which is called loadings and can be obtained from the previously mentioned Power method.

The $\mathbf{t}_i$ vectors form an orthogonal set ($\mathbf{t}_i'\mathbf{t}_j = 0$ for $i \neq j$), while the $\mathbf{p}_i$ are orthonormal ($\mathbf{p}_i'\mathbf{p}_j = 0$ for $i \neq j$, $\mathbf{p}_i'\mathbf{p}_j = 1$ for $i = j$). Eq. (4.10) can be rewritten as

$$\mathbf{X} = \mathbf{T}\mathbf{P}' + \mathbf{E} \tag{4.11}$$

Note that for **X** and any $\mathbf{t}_i$-$\mathbf{p}_i$, the following equation holds:

$$\mathbf{t}_i = \mathbf{X}\mathbf{p}_i \tag{4.12}$$

This means that the score vector $\mathbf{t}_i$ is the linear combination of the original **X** data defined by $\mathbf{p}_i$. $\mathbf{t}_i$ can also be viewed as the projection of **X** onto $\mathbf{p}_i$. Here $\mathbf{p}_i$ represents the new coordinate axis (or direction) for the space spanned by **X**. Since $\mathbf{t}_i$ is orthogonal to each other, they are new uncorrelated variables of the original data in **X**. Therefore, PCA transforms n correlated variables into k uncorrelated variables. Because the original data is mean-centered and standard deviation scaled, the elements of the loadings $\mathbf{p}_i$ will range from –1 to 1 with high absolute values corresponding to high correlations. The first eigenvector describes the direction of the greatest variation in **X**; the second describes the next dominant direction of variation and so on. From Eq. (4.4), there exits a useful property:

$$\mathrm{Tr}(\mathbf{S}) = \mathrm{Tr}(\mathbf{L}) \tag{4.13}$$

that is, the sum of the original variances is equal to the sum of the characteristic roots. From this, it can be concluded that the characteristic roots, or the variance of the principal components, can be treated as variance components. The ratio of each characteristic root to the total indicates the proportion of the total variability accounted for by each principal component (PC). Since the characteristic roots are sample estimates, these proportions are also sample estimates.

Consequently, the PCs are arranged in a descending order according to the corresponding l_i. l_i can be viewed as a measure of the amount of variance described by the principal components. Because PCs are in the descending order of l_i, the first principal component, PC1, captures the greatest amount of variation in the data. Each

subsequent PC captures the greatest possible amount of variance remaining after subtracting $\mathbf{t}_i\mathbf{p}_i'$ from $\mathbf{X}$.

Autoscaling is needed where the variables have different units or in different scales, e.g. temperature, pressure and concentration. A variation in temperature between 1000 and 1100 °C is much greater than a variation in concentration between 0.01 and 0.1. However, the effect of each of these variables on the system may be very similar and the information content of the temperature data is not inherently greater. For this reason, autoscaling the data is required. This will put all variables on an equal bases in the analysis.

Using PCA, one can represent an n-variable data set in $k < n$ dimensions. This is its data compression capability. For a larger k, a better fit PCA model can be built but calculation may be difficult. If $k = n$, then all information will be preserved in the model while the analysis becomes very complicated. There is an optimal value of k for a given data. In this research, the purpose of PCA performance is not to build a PCA model but to obtain informative variables. Based on the correlation coefficients between PCs and the target property, the value of k corresponding to the most informative variables can be chosen. This will be discussed in detail later.

4.5 Geometric interpretation of PCA

Consider a covariance matrix computed from a bivariate data set, $\mathbf{x}_1$ and $\mathbf{x}_2$ as below:

$$\mathbf{S} = \begin{bmatrix} 20.28 & 15.58 \\ 15.58 & 24.06 \end{bmatrix} \tag{4.14}$$

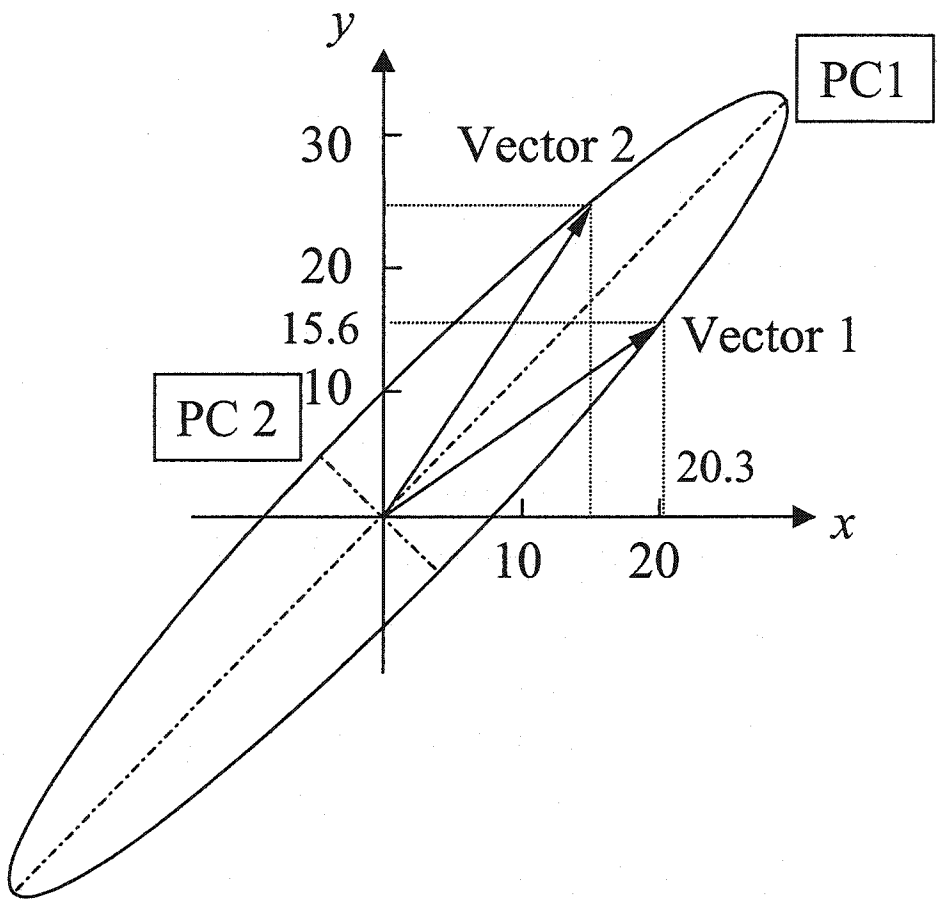

Fig. 4.2 The covariance plotted as vectors 1 and 2. The ellipse major axis is the first eigenvector (of unit length) multiplied by the corresponding eigenvalue. The minor axis represents the second eigenvector/value.

We see that the covariance is symmetric about the diagonal, $\mathbf{x}_1$ and $\mathbf{x}_2$. If we extract the eigenvectors and eigenvalues from this covariance matrix, we have a new set of basic functions to represent the data where the covariance matrix was derived. We can consider the columns of the covariance matrix **S** in Eq. (4.14) as vectors and plot them with an ellipse representing the orientation of the eigenvectors and the magnitudes of the eigenvalues (See Fig. 4.2). Vectors 1 and 2 represent a basis for this data set, but they are not independent. Because the eigenvectors are always orthonormal they are always

independent and more efficient to represent the original data. If, in Fig. 4.2 we project vector 1 (the first vector formed from the covariance matrix) back onto the major and minor axes of the ellipse (the first and second eigenvectors), we get the more efficient representation. Most of the information is loaded onto the first principal component and this is true for each individual sample as well. We call these more efficient coordinates the principal component axis.

4.6 Summary

Principal component analysis (PCA) applied in this research is to meet the challenge of reducing the system information redundancy and extracting information most relevant to the parameter to be determined. PCA is a multivariate technique of chemometrics by which a number of related variables are transformed to a set of uncorrelated variables called principal components (PC). Each component is a linear combination of several original variables. Since a PC combines the contributions of several original variables, it may be more sensitive to a particular event than an individual variable. The PC that captures the largest amount of variation in the data is called the 1^{st} component, and the PC that captures the 2^{nd} largest amount of data variation is called the 2^{nd} component, etc. By studying the correlation between each PC and the parameter to be determined, one can choose to use only those PCs most related to the parameter of interest. Through this technique, the information redundancy can be reduced and useful information can be extracted efficiently.

Chapter 5 Artificial Neural Networks

5.1 Background

Artificial neural network (ANN) is a network that consists of a large amount of small elements ("neurons") to local process input data and to convey data between the neurons [2]. Using learning algorithms, the connection weights can be adjusted according to the input data set during the training period, and hence, the neural network has the capability of "learning". More specifically, training is effected by presenting the network with the known inputs and outputs and modifying the connection weights between individual neurons, following certain back- propagation algorithm [60]. The training process continues until the output of the network matches the desired outputs to a stated degree of accuracy. The trained ANNs may then be exposed to unknown inputs (i.e. spectra) to provide best fit to the outputs.

According to Sarle [63], "Neural networks (NNs) are especially useful for classification and function approximation/mapping problems which are tolerant of some imprecision, which have lots of training data available, but to which hard and fast rules (such as those that might be used in an expert system) cannot easily be applied." Therefore, neural networks can be applied to process large amount of data and characterize complex systems difficult to model by standard procedures. For this reason, we use ANN to process the collected data set of polymer extrusion process.

The concept of ANN was first introduced by McCulloch and Pitts [43]. Considerable interest in this field has been renewed since early 1980s due to the

introduction of Hopfield's energy approach [26] and the backpropagation learning algorithm proposed by Werbs [77]. From its historical account of developments [1], the field of artificial neural networks is typically an interdisciplinary area of research. A thorough study of ANNs requires some basic knowledge from neurophysiology, control theory, mathematics, statistics, decision making, and distributed computing. In this research, only needed background is reviewed. For more detailed information or features of ANNs, the reader is referred to [5,30,40,44,57].

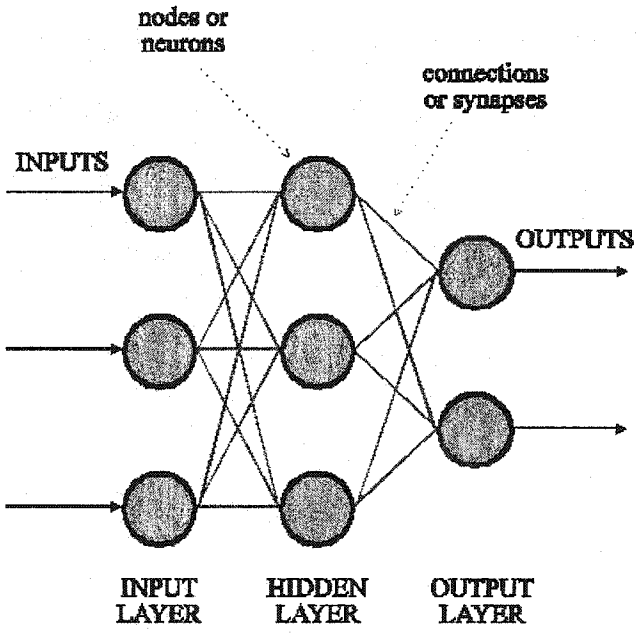

Fig. 5.1 Basic scheme of artificial neural network

In its most general form, an artificial neural network is a machine that is designed to model the way in which the brain performs a particular task. The network is usually implemented using electronic components or is simulated in software on a digital computer. To achieve good performance, ANN employ a massive interconnection of

simple computing cells referred to as "neurons" or "processing units." Then, the definition of an artificial neural network is as follows [44]:

A neural networks is a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the human brain in two respects:

1. Knowledge is acquired by the network from its environment through a learning process.
2. Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

ANNs employ a large set of processing nodes and the data are acted upon in parallel. As shown in Fig. 5.1, almost all ANNs employ input nodes for data entry, output nodes for network classification, and intermediate nodes in hidden layers for data analysis. Based on the node arrangement, the interconnections and the training algorithm applied, there are many differing types of ANNs. In general, ANNs are trained using either supervised or unsupervised learning algorithms. In the supervised case, the network is presented with pre-selected signals representing various classes and is trained to recognize them. The main difference between supervised and unsupervised learning lies in that the latter can select its own training set. Although there are many theoretical studies on both types of ANNs, a layered-node arrangement of supervised category known as back-propagation (BP) net is most popular [44]. It utilizes the network output for the training based on certain criterion. Each processing node or neuron in a given layer except the input layer makes a decision on the basis of the decision boundary it is

trained to recognize when presented with a set of data. Its decision is then used by the succeeding layer until the final output neuron layer is reached. Hence, the output node layer classifies the unknown signal into one of the predefined classes.

Classification of signals into distinct categories is made on the basis of a pre-selected signal set defining each class. This pre-selected signal set, known as the training set, contains the data related to each separate class and is used to train the ANNs.

For supervised learning networks, the selection of training signal is very important. Actually, the accuracy of the obtained classification using back-propagation depends heavily on the choice of signals applied as the training data set. If the training signals are not truly representative, the classification will not be accurate. For example, a network trained by large flaw size signals may incorrectly classify signals with small flaws as signals with no defects. For our research, more serious problem is the heavy data correlation making data processing very difficult. Moreover, large mount of original data input to the network increases the calculation time and the complexity of the network. Since PCA is able to reduce the dimensionality of the original data while preserves most of the relevant information in the new low-dimension space, it was used to pre-process the data to improve the network. To determinate the index of dispersion, there are two major steps in this research. First step is to use PCA to select the most informative variables to construct the feature data sets. Secondly, feed these feature data into a neural network to determine the dispersion index.

There are many publications on applications of ANNs in chemistry [21,53,79], pharmaceutics [80], chemical engineering [12,81], and face recognition [37]. The

combination of ANNs and other techniques, such as pyrolysis mass spectrometry [23], ultrasonics [7,14,15,25] were reported. Methods combining PCA and ANNs have been developed for pattern recognition [35,41,42], stellar spectrum analysis [64,66].

5.2 Notations

The procedure used to perform the learning process is called a learning algorithm. It is to modify the synaptic weights of the network in an orderly fashion to attain a desired design objective.

We present notations to be used in the rest of this chapter:

- Indices i, j, and k - refer to different neurons in the network.
- Iteration (time step) n - the nth training pattern (example) presented to the network.
- $E(n)$ - instantaneous sum of error squares or error energy at iteration n.
- $e_j(n)$ - error signal at the output of neuron j at iteration n.
- $d_j(n)$ - desired response for neuron j and is used to compute $e_j(n)$.
- $y_j(n)$ – output, function signal of neuron j at iteration n.
- $w_{ji}(n)$ – the synaptic weight connecting the output of neuron i to the input of neuron j at iteration n.
- $\Delta w_{ji}(n)$ – The correction applied to the weight $w_{ji}(n)$ at iteration n.
- $v_j(n)$ – induced local field (i.e., weighted sum of all synaptic inputs plus bias) of neuron j at iteration n.
- $\varphi(\cdot)$ – activation function describing the input-output functional relationship of the nonlinearity associated with neuron j.

- $\delta_j(n)$ – local gradient for neuron j.
- b_j – bias of neuron j.
- $x_i(n)$ – ith element of the overall input vector.
- $o_k(n)$ – kth element of the overall output vector.
- η – learning rate parameter.
- m_l – size (i.e., number of nodes) in layer l of the multilayer perceptron, $l = 0,1,\ldots,L$, where L is the "depth" of the network.
- m_0 – size of the input layer, m_1 denotes the size of the first hidden layer.
- m_L – size of the output layer. The notation $m_L = M$ is also used.

In using the above notations, we make the following remarks:

- Indices i, j, and k are arranged as: with signals propagating through the network from left to right, neuron j lies in a layer to the right of neuron i, and neuron k lies in a layer to the right of neuron j when neuron j is a hidden unit.
- The average of $E(n)$ over all values of n (i.e., the entire training set) yields the average error energy E_{av}.
- $d_j(n)$ is used to compute $e_j(n)$.
- $v_j(n)$ constitutes the signal applied to the activation function associated with neuron j.
- the effect of b_j is represented by a synapse of weight $w_{j0} = b_j$ connected to a fixed input equal to +1.

5.3 Learning Process

An ANN has a set of sensory units (source nodes) that constitute the input layer, one or more hidden layers of computation nodes, and an output layer of computation nodes. The input signal propagates through the network in a forward direction, on a layer-by-layer basis. The most attractive property of an artificial neural network is the ability of the network to learn from its environment, and to improve its performance through learning. Through an interactive process of adjustments applied to its synaptic weights and bias levels, an ANN tries to learn about its environment, and ideally, becomes more knowledgeable about its environment after each iteration of the learning process. Generally, these neural networks are referred to as multilayer perceptrons (MLPs). Multilayer perceptrons have demonstrated their capabilities by successfully solving some difficult and diverse problems through training them in a supervised manner with the error back-propagation algorithm. This algorithm is based on an error-correction learning rule. Error back-propagation learning consists of two phases through different layers of the network: a forward phase and a backward phase. In the forward phase, an activity pattern (input vector) is applied to the sensory nodes of the network, and its effect propagates through the network layer by layer. Finally, a set of outputs is produced as the actual response of the network. In the forward phase, the synaptic weights of the networks are all fixed. In the backward phase, on the other hand, the synaptic weights are all adjusted according to an error-correction rule. The actual response of the network is subtracted from a desired response to produce an error signal.

This error signal is then propagated backward through the network, against the direction of synaptic connections – hence the name "error back-propagation." The synaptic weights are adjusted to make the actual response of the network move closer to the desired response in a statistical sense. The error back-propagation algorithm is also referred to in the literature as the back-propagation algorithm, or simply back-prop. The learning process performed with the algorithm is called back-propagation learning.

5.4 Back-propagation algorithm

The back-propagation algorithm is used to adjust the synaptic weights during the learning process. The error signal at the output of neuron j at iteration n is defined by:

$$e_j(n) = d_j(n) - y_j(n), \text{ neuron } j \text{ is an output node.} \tag{5.1}$$

We define the instantaneous value of the error energy for neuron j as $\frac{1}{2}e_j^2(n)$. Correspondingly, the instantaneous value $E(n)$ of the total error energy is obtained by summing $\frac{1}{2}e_j^2(n)$ over all neurons in the output layer. These are the only "visible" neurons for which error signals can be directly calculated. $E(n)$ can be written as:

$$E(n) = \frac{1}{2}\sum_{j \in C} e_j^2(n) \tag{5.2}$$

where the set C includes all the neurons in the output layer of the network. Let N denote the total number of patterns (examples) contained in the training set. The average squared error energy is obtained by summing $E(n)$ over all n and then normalized with respect to N, as shown by:

$$E_{av} = \frac{1}{N} \sum_{n=1}^{N} E(n) \tag{5.3}$$

The instantaneous error energy $E(n)$, and therefore the average error energy E_{av}, is a function of the free parameters (i.e., synaptic weights and bias levels) of the network. For a given training set, E_{av} represents the cost function or performance index as a measure of the learning performance. The objective of the learning process is to adjust the free parameters of the network to minimize E_{av}. To do this minimization, an approximation approach can be utilized. A simple method of training is to update the weights on a pattern-by-pattern basis until one complete presentation of the entire training set is conducted. The adjustments to the weights are made in accordance with respective errors computed for each pattern presented to the network. The arithmetic average of these individual weight changes over the training set is therefore an estimate of the true change that would result from modifying the weights based on minimizing the cost function E_{av} over the entire training set.

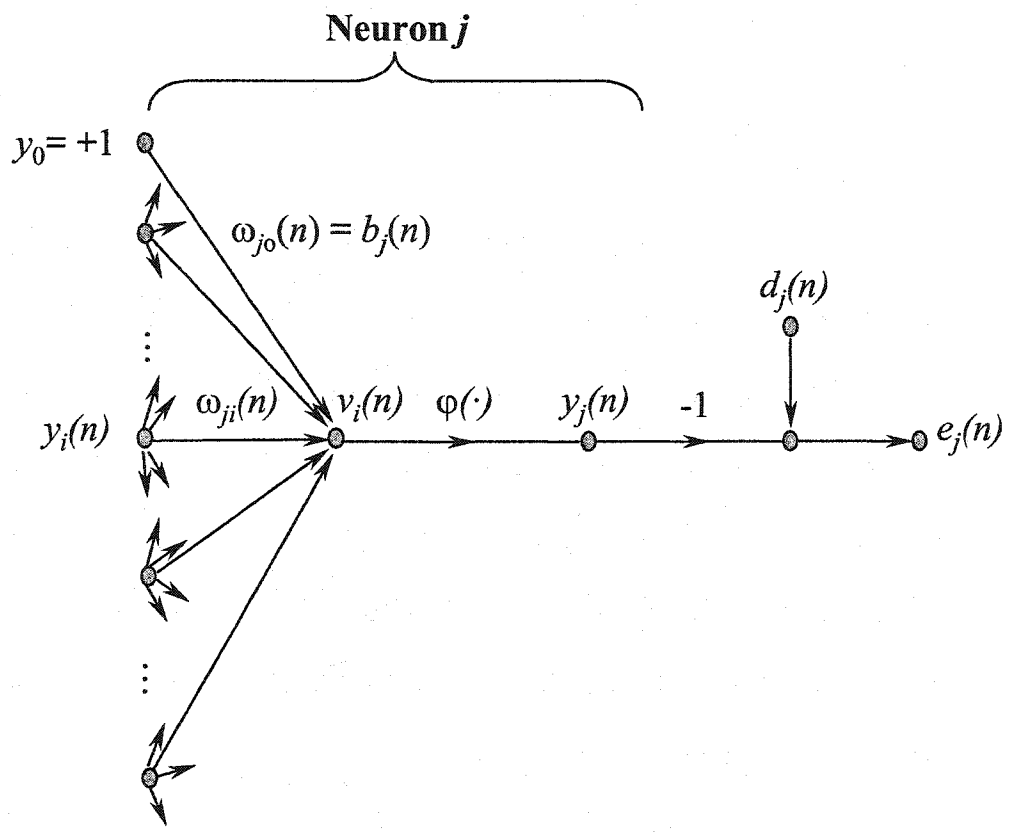

Fig. 5.2 Signal-flow graph highlighting the details of output neuron *j*

Consider the signal-flow of an output neuron (see Fig. 5.2), where neuron *j* is fed by a set of function signals produced by a layer of neurons to its left. The induced local field $v_j(n)$ produced at the input of the activation function associated with neuron *j* is therefore:

$$v_j(n) = \sum_{j=0}^{m} w_{ji}(n) y_i(n) \tag{5.4}$$

where *m* is the total number of inputs, excluding the bias, applied to neuron *j*. The synaptic weight w_{j0} corresponding to the fixed input $y_0 = +1$, equals to the bias b_j applied to neuron *j*. Hence the function signal $y_j(n)$ at the output of neuron *j* at iteration *n* is:

$$y_j(n) = \varphi_j(v_j(n)) \tag{5.5}$$

Then the back-propagation algorithm applies a correction $\Delta w_{ji}(n)$ to the synaptic weight $w_{ji}(n)$, which is proportional to the partial derivative $\partial E(n)/\partial w_{ji}(n)$. It can be written as:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \tag{5.6}$$

The partial derivative $\partial E(n)/\partial w_{ji}(n)$ represents a sensitivity factor, determines the direction of search in the weight space for the synaptic weight w_{ji}. Differentiating both sides of Eq. (5.2) with respect to $e_j(n)$, we get

$$\frac{\partial E(n)}{\partial e_j(n)} = e_j(n) \tag{5.7}$$

Differentiating both sides of Eq. (5.1) with respect to $v_j(n)$, we get

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \tag{5.8}$$

Differentiating Eq. (5.5) with respect to $v_j(n)$, we get

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi_j'(v_j(n)) \tag{5.9}$$

where the prime of $\varphi_j(\cdot)$ denotes differentiation with respect to its argument. Finally, differentiating Eq. (5.4) with respect to $w_{ji}(n)$ yields

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_i(n) \tag{5.10}$$

Put Eqs.(5.7) to (5.10) in (5.6), we have:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n)\varphi'_j(v_j(n))y_i(n) \tag{5.11}$$

The correction $\Delta w_{ji}(n)$ applied to $w_{ji}(n)$ is defined by the delta rule:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \tag{5.12}$$

where η is the learning-rate parameter of the back-propagation algorithm. The use of the minus sign in Eq.(5.12) accounts for gradient descent in seeking a direction for weight change to reduce the value of $E(n)$. Put Eq.(5.11) in (5.12), we have

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_j(n) \tag{5.13}$$

where the local gradient $\delta_j(n)$ is defined by

$$\begin{aligned} \delta_j(n) &= -\frac{\partial E(n)}{\partial v_j(n)} \\ &= -\frac{\partial E(n)}{\partial e_{j(n)}} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \\ &= e_j(n)\varphi'_j(v_j(n)) \end{aligned} \tag{5.14}$$

The local gradient points to required changes in the synaptic weights. According to Eq.(5.14), the local gradient $\delta_j(n)$ for output neuron j is equal to the product of the error signal $e_j(n)$ and derivative $\varphi'_j(v_j(n))$ of the activation function. From Eqs.(5.13) and

(5.14) we note that a key factor involved in the calculation of the weight adjustment $\Delta w_{ji}(n)$ is the error signal $e_j(n)$ in the output of neuron j. It is treated differently for neuron j at different layers.

5.4.1 Output layer

When neuron j is located in the output layer of the network, it is supplied with a desired response of its own. The error signal $e_j(n)$ associated with this neuron is computed using Eq.(5.1). Having determined $e_j(n)$, it is straightforward to compute the local gradient $\delta_j(n)$ using Eq.(5.14).

5.4.2 Hidden layer

When neuron j is located in a hidden layer of the network, there is no specified desired response for that neuron. Accordingly, the error signals for a hidden neuron will be determined recursively in terms of the error signals of all the neurons to which that hidden neuron is directly connected. This makes the development of the back-propagation algorithm complicated. Based on Eqs.(5.9) and (5.14), we may redefine the local gradient $\delta_j(n)$ for hidden neuron j as

$$\begin{aligned}\delta_j(n) &= -\frac{\partial E(n)}{\partial y_j(n)}\frac{\partial y_j(n)}{\partial v_j(n)} \\ &= -\frac{\partial E(n)}{\partial y_j(n)}\varphi_j'(v_j(n)),\end{aligned} \tag{5.15}$$

To calculate the partial derivative $\partial E(n)/\partial y_j(n)$, one proceeds as follows. From Fig.5.2 one has:

$$E(n) = \frac{1}{2}\sum_{k \in C} e_k^2(n), \quad \text{neuron } k \text{ is an output node} \tag{5.16}$$

which is the same as Eq.(5.2) with index k in place of index j. Differentiating Eq.(5.16) with respect to function signal $y_j(n)$, we get

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial y_j(n)} \tag{5.17}$$

Next we use the chain rule for the partial derivative $\partial e_k(n)/\partial y_j(n)$, and rewrite Eq.(5.17) to:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial y_j(n)} \tag{5.18}$$

From Fig.5.2, it follows

$$\begin{aligned} e_k(n) &= d_k(n) - y_k(n) \\ &= d_k(n) - \varphi_k(v_k(n)), \end{aligned} \quad \text{neuron } k \text{ is an output node} \tag{5.19}$$

Hence

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi_k'(v_k(n)) \tag{5.20}$$

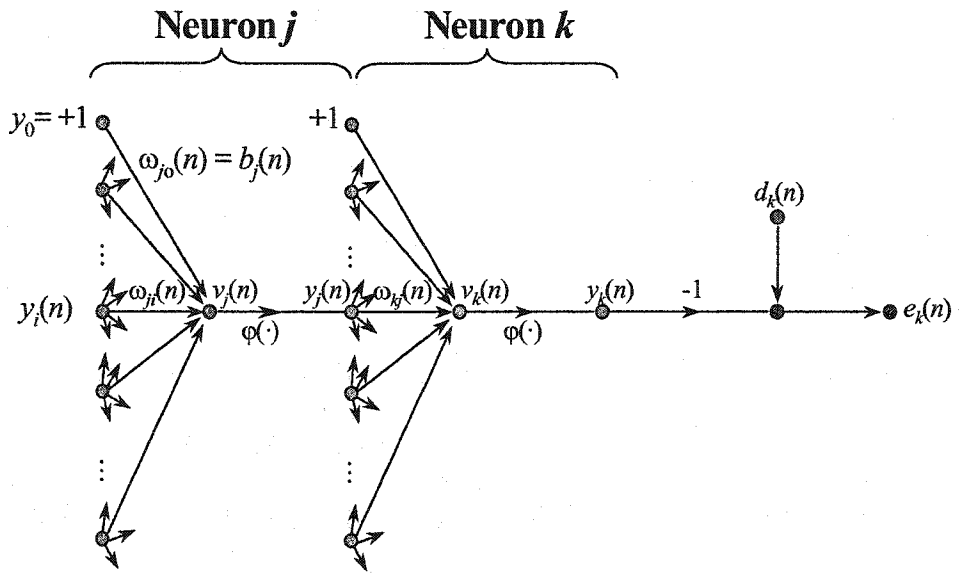

Fig. 5.3 Signal-flow graph highlighting the details of output neuron k connected to hidden neuron j

We also note from Fig.5.2 that, for neuron k, the induced local field is

$$v_k(n) = \sum_{j=0}^{m} w_{kj}(n) y_j(n) \tag{5.21}$$

where m is the total number of inputs applied to neuron k. Here again, the synaptic weight $w_{k0}(n)$ is equal to the bias $b_k(n)$ applied to neuron k, and the corresponding input is fixed at the value +1. Differentiating Eq.(5.21) with respect to $y_j(n)$ yields

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj}(n) \tag{5.22}$$

Using Eqs.(5.20) and (5.22) in (5.18), one gets the desired partial derivative:

$$\begin{aligned} \frac{\partial E(n)}{\partial y_j(n)} &= -\sum_k e_k(n) \varphi_k'(v_k(n)) w_{kj}(n) \\ &= -\sum_k \delta_k(n) w_{kj}(n) \end{aligned} \tag{5.23}$$

where in the second part of Eq.(5.23) we have used the definition of the local gradient $\delta_k(n)$ given in Eq.(5.14) with index k replacing j.

Finally, using Eq.(5.23) in (5.15), we get the back-propagation formula for the local gradient $\delta_j(n)$ as described:

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n), \tag{5.24}$$

where neuron j is in hidden layer. The factor $\varphi'_j(v_j(n))$ involved in the computation of the local gradient $\delta_j(n)$ in Eq. (5.24) depends solely on the activation function of the hidden neuron j. The remaining factor involved in this computation, namely the summation over k, depends on $\delta_k(n)$ and $w_{kj}(n)$. To calculate $\delta_k(n)$ requires knowledge of the error signals $e_k(n)$, for all neurons that lie in the layer to the immediate right of hidden neuron j and directly connected to it (See Fig.5.2). $w_{kj}(n)$ consists of the synaptic weights associated with these connections.

In summary, the correlation $\Delta w_{ji}(n)$ is first applied to the synaptic weight connecting neuron i to neuron j is defined by the delta rule:

$$\begin{pmatrix} Weight \\ correction \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} learning- \\ rate\ parameter \\ \eta \end{pmatrix} \cdot \begin{pmatrix} local \\ gradient \\ \delta_j(n) \end{pmatrix} \cdot \begin{pmatrix} input\ signal \\ of\ neuron\ j \\ y_i(n) \end{pmatrix} \tag{5.25}$$

Then calculation of local gradient $\delta_j(n)$ depends on whether neuron j is an output node or a hidden node:

1. If neuron j is an output node, $\delta_j(n)$ equals the product of the derivative $\varphi'_j(v_j(n))$ and the error signal $e_j(n)$, as shown in Eq.(5.14).

2. If neuron j is a hidden node, $\delta_j(n)$ equals the product of the associated derivative $\varphi'_j(v_j(n))$ and the weighted sum of the δs computed for the neurons in the next hidden or output layer that are connected to neuron j as shown in Eq.(5.24).

5.5 Two computational phases

There are two distinct phases in the application of the back-propagation algorithm. The first phase is the forward phase, and the second phase is the backward phase. In the forward phase the synaptic weights remain unaltered throughout the network, and the function signals of the network are computed on a neuron-by-neuron basis. The function signal appearing at the output of neuron j is computed by

$$y_j(n) = \varphi(v_j(n)) \tag{5.26}$$

where $v_j(n)$ is the induced local field of neuron j, defined by

$$v_j(n) = \sum_{i=0}^{m} w_{ji}(n) y_i(n) \tag{5.27}$$

where m is the total number of inputs (excluding the bias) applied to neuron j, and $w_{ji}(n)$ is the synaptic weight connecting neuron i to neuron j. $y_i(n)$ is the input signal of neuron j or equivalently, the function signal of the output of neuron i. If neuron j is in the first hidden layer of the network, then $m = m_0$. The index i refers to the ith input of the network, and

$$y_i(n) = x_i(n) \tag{5.28}$$

where $x_i(n)$ is the ith element of the input vector. If neuron j is in the output layer of the network, then $m = m_L$ and j refers to the jth output of the network, we then have:

$$y_j(n) = o_j(n) \tag{5.29}$$

where $o_j(n)$ is the jth element of the output. This output is compared with the desired response $d_j(n)$, obtaining the error signal $e_j(n)$ for the jth output neuron. Thus the forward phase of computation begins at the first hidden layer and terminates at the output layer by computing the error signal for each neuron of this layer.

The backward phase starts from the output layer through the network, layer by layer, and recursively computes the local gradient δ_j for each neuron. This recursive process permits the synaptic weights of the network to undergo changes in accordance with the delta rule of Eq.(5.25). For a neuron located in the output layer, the δ_j equals to the error signal of that neuron multiplied by the first derivative of its nonlinear function. We use Eq.(5.25) to compute the changes of the weights of all the connections feeding into the output layer. Given the δ_j for the neurons in the output layer, we use Eq.(5.24) to compute the δ_j for all the neurons in the penultimate layer and the changes to the weights of all the connections feeding into it. The recursive computation is continued, layer by layer, by propagating the changes to all the synaptic weights in the network.

5.6 Activation function

The computation of δ_j for each neuron of the multilayer perceptron requires the derivative of the activation function $\varphi(\cdot)$ associated with that neuron. For this derivative to exist, the function $\varphi(\cdot)$ must be continuous and differentiable. A commonly used continuously

differentiable nonlinear activation function in multilayer perceptrons is the sigmoidal nonlinear function. It has two forms:

1. *Logistic Function.* This form of sigmoidal nonlinearity in its general form is given by [40]

$$\varphi_j(v_j(n)) = \frac{1}{1+\exp(-av_j(n))}, a > 0 \, and \, -\infty < v_j(n) < \infty \tag{5.30}$$

where $v_j(n)$ is the induced local field of neuron j. Using this function, the amplitude of the output is in the range of $0 \leq y_j(n) \leq 1$. Differentiating Eq.(5.30) with respect to $v_j(n)$, we get

$$\varphi_j'(v_j(n)) = \frac{a\exp(-av_j(n))}{[1+\exp(-av_j(n))]^2} \tag{5.31}$$

With $y_j(n) = \varphi_j(v_j(n))$, we can eliminate the exponential term $\exp(-a \cdot v_j(n))$ from Eq.(5.31) and express $\varphi_j'(v_j(n))$ by

$$\varphi_j'(v_j(n)) = ay_j(n)[1-y_j(n)] \tag{5.32}$$

For a neuron j located in the output layer, $y_j(n) = o_j(n)$, we can express the local gradient of neuron j by

$$\begin{aligned} \delta_j(n) &= e_j(n)\varphi_j'(v_j(n)) \\ &= a[d_j(n)-o_j(n)]o_j(n)[1-o_j(n)] \end{aligned} \tag{5.33}$$

where $o_j(n)$ is the function signal at the output of neuron j and $d_j(n)$ is the desired response. For an arbitrary hidden neuron j, we can express the local gradient by

$$\delta_j(n) = \varphi_j'(v_j(n))\sum_k \delta_j(n) w_{kj}(n)$$
$$= a y_j(n)[1 - y_j(n)]\sum_k \delta_k(n) w_{kj}(n), \text{ neuron } j \text{ is hidden} \qquad (5.34)$$

From Eq.(5.32), $\varphi_j'(v_j(n))$ attains its maximum value at $y_j(n) = 0.5$, and its minimum value at $y_j(n) = 0$, or $y_j(n) = 1.0$. Since the amount of change in a synaptic weight of the network is proportional to $\varphi_j'(v_j(n))$. For a sigmoid activation function, the synaptic weights change most rapidly when the function signals are in their midrange.

2. *Hyperbolic tangent function.* Another commonly used form of sigmoidal nonlinear function is the hyperbolic tangent function. The general form of the function is given by

$$\varphi_j(v_j(n)) = a\tanh(b v_j(n)), \quad (a,b) > 0 \qquad (5.35)$$

where a and b are constants. The hyperbolic tangent function is a rescaled and biased logistic function. Its derivative with respect to $v_j(n)$ is given by

$$\begin{aligned} \varphi_j'(v_j(n)) &= ab\,\mathrm{sec}\,h^2(b v_j(n)) \\ &= ab(1 - \tanh^2(b v_j(n))) \\ &= \frac{b}{a}[a - y_j(n)][a + y_j(n)] \end{aligned} \qquad (5.36)$$

For a neuron j located in the output layer, the local gradient is

$$\begin{aligned} \delta_j(n) &= e_j(n)\varphi_j'(v_j(n)) \\ &= \frac{b}{a}[d_j(n) - o_j(n)][a - o_j(n)][a + o_j(n)] \end{aligned} \qquad (5.37)$$

For a neuron j in a hidden layer, we have

$$\begin{aligned}\delta_j(n) &= \varphi_j'(v_j(n))\sum_k \delta_k(n) w_{kj}(n) \\ &= \frac{b}{a}[a - y_j(n)][a + y_j(n)]\sum_k \delta_k(n) w_{kj}(n)\end{aligned} \tag{5.38}$$

Using Eqs.(5.33) and (5.34) for the logistic function and Eqs.(5.37) and (5.38) for the hyperbolic tangent function, we can calculate the local gradient δ_j without requiring explicit knowledge of the activation function.

5.7 Sequential and batch modes of training

In practical applications of the back-propagation algorithm, learning results from a set of training examples to the multilayer perceptron. One complete presentation of the entire training set is an epoch. The learning process is maintained on an epoch-by-epoch basis until the synaptic weights and bias levels of the network stabilize and the average squared error over the entire training set converges to some minimum value. One needs to randomize the order of presentation of training examples from one epoch to the next. This randomization makes the search in weight space random over the learning cycles and to avoid the possibility of cycles in the evolution of the synaptic weight vectors. For a given training set, back-propagation learning may thus proceed in one of the following two basic ways:

1. Sequential Mode. The sequential mode of back-propagation learning is also referred to as on-line, pattern, or stochastic mode. In this mode of operation weight updating is performed after the presentation of each training example. For an epoch consisting of N training examples arranged in the order of ($\mathbf{x}(1)$, $\mathbf{d}(1)$), ..., ($\mathbf{x}(N)$, $\mathbf{d}(N)$). The first example pair ($\mathbf{x}(1)$, $\mathbf{d}(1)$) in the epoch is presented, and the sequence of forward

and backward computations described previously is performed, resulting in certain adjustments to the synaptic weights and bias levels of the network. Then the second example pare (**x**(2), **d**(2)) in the epoch is presented, and the sequence of forward and backward computations is repeated, resulting in further adjustments to the synaptic weights and bias levels. This process continues until the last example pair (**x**(*N*), **d**(*N*)) in the epoch is presented.

2. Batch Mode. In the batch mode of back-propagation learning, weight updating is performed after presentation of all the training examples that constitute an epoch. For a particular epoch, we define the cost function as the average squared error in Eqs.(5.2) and (5.3) in the form of:

$$E_{av} = \frac{1}{2N} \sum_{n=1}^{N} \sum_{j \in C} e_j^2(n) \tag{5.39}$$

where the error signal $e_j(n)$ pertains to output neuron j for training example n defined in Eq.(5.1). Error $e_j(n)$ is the difference between $d_j(n)$ and $y_j(n)$, the jth element of the desired response vector $\mathbf{d}(n)$ and the corresponding value of the network output, respectively. In Eq.(5.39) the inner summation with respect to j is performed over all the neurons in the output layer of the network, and the outer summation with respect to n is performed over the entire training in the epoch. For a learning-rate parameter η, the adjustment applied to synaptic weight w_{ji}, connecting neuron i and neuron j, is defined by the delta rule

$$\Delta w_{ji} = -\eta \frac{\partial E_{av}}{\partial w_{ji}} = -\frac{\eta}{N}\sum_{n=1}^{N} e_j(n)\frac{\partial e_j(n)}{\partial w_{ji}} \tag{5.40}$$

To calculate the partial derivative $\partial e_j(n)/\partial w_{ji}$ we proceed in the same way as before. Using Eq.(5.40), the weight adjustment Δw_{ji} is made after the entire training set has been presented to the network.

5.8 Stopping criteria

The back-propagation algorithm in general may not to converge, and there are no widely accepted criteria for stopping the iteration. Some reasonable criteria, each with its own practical merit, may be used to terminate the weight adjustments. To formulate such a criterion, it is logical to think in terms of the unique properties of a local or global minimum of the error surface. Let the weight vector **w*** denote a minimum, be it local or global. Here, a vector **w*** is said to be a local minimum of an input-output function F if it is no worse than its neighbors, that is, if there exists an ε such that [6]

$$F(\mathbf{w}^*) \leq F(\mathbf{w}) \qquad \text{for all } \mathbf{w} \text{ with } \|\mathbf{w} - \mathbf{w}^*\| < \varepsilon$$

The vector **w*** is said to be a global minimum of the function F if it is no worse than all other vectors; that is,

$$F(\mathbf{w}^*) \leq F(\mathbf{w}) \qquad \text{for all } \mathbf{w} \in \Re^n$$

where n is the dimension of **w**.

A necessary condition for **w*** to be a minimum is that the gradient vector **g(w)** of the error surface with respect to the weight vector **w** be zero at **w** = **w***. The back-propagation algorithm is considered to have converged when the Euclidean norm of the gradient vector reaches a sufficiently small threshold. The drawback of this convergence criterion is that, for successful trials, learning times may be long. Also, it requires the computation of the gradient vector **g(w)**.

We can consider a different criterion of convergence. The back-propagation algorithm is considered to have converged when the absolute rate of change in the average squared error per epoch is sufficiently small. The rate of change in the average squared error is typically considered to be small if it lies in the range of 0.1% to 1% per epoch. Sometimes even 0.01% per epoch is used.

5.9 Rate of learning

The back-propagation algorithm finds the approximate trajectory in the weight space using the steepest descent method. If the learning-rate parameter η is small, the process will be smooth but the computation may be very slow. If we use a large learning-rate parameter η, it may result in large changes in the synaptic weights and the network may become unstable. A simple method to increase the rate of learning and to avoid the danger of instability is to modify the delta rule in Eq.(5.13) by including a momentum term. The generalized delta rule can be used as discussed in [59]:

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \tag{5.41}$$

where α, a positive number, is the momentum constant. It controls the feedback loop acting around $\Delta w_{ji}(n)$. When $\alpha = 0$, Eq.(5.41) becomes the delta rule of Eq.(5.13).

We rewrite Eq. (5.41) as a series:

$$\Delta w_{ji}(n) = \eta[\alpha^{n-1}\delta_j(1)y_i(1) + \alpha^{n-2}\delta_j(2)y_i(2) + \cdots + \alpha\delta_j(n-1)y_i(n-1) + \delta_j(n)y_i(n)] \tag{5.41a}$$

Let an index t go from the initial time 0 to the current iteration n. Then we can further rewrite Eq. (5.41a) as:

$$\Delta w_{ji}(n) = \eta \sum_{t=0}^{n} \alpha^{n-t}\delta_j(t)y_i(t) \tag{5.42}$$

It represents a time series of length n+1. From Eqs. (5.11) and (5.14) we note $\delta_j(n)y_j(n) = -\partial E(n)/\partial w_{ji}(n)$. We rewrite Eq.(5.42)

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^{n} \alpha^{n-t} \frac{\partial E(t)}{\partial w_{ji}(t)} \tag{5.43}$$

From Eq.(5.43), we can use an α with $|\alpha|<1$ so that weight adjustment process will converge.

5.10 Summary

In this chapter, several important concepts were introduced in sequences, such as back-propagation algorithm, forward/backward computation phases, activation function, training modes, stopping criteria, and learning rates. Learning is the core of artificial

neural networks to allow the network to learn from its environment and improve its performance. During the process, the input signal flows through the network neurons layer-by-layer and in a forward direction. The error signal is then propagated backward through the network. The synaptic weights are adjusted to make the actual response of the network move closer to the desired response.

The model of each neuron in the ANN includes a nonlinear activation function. The presence of nonlinearities is important. Otherwise, the ANN will be reduced to a single-layer perceptron without proper network functionality. The use of the logistic function is biologically motivated. It attempts to account for the refractory phase of real neurons. The artificial neural network contains one or more layers of hidden neurons that are not part of the input or output of the network. These hidden neurons enable the network to learn complex tasks by extracting progressively more meaningful features from the input patterns. From the combination of those characteristics and the ability to learn from experience through training, the multilayer perceptron derives it computing power. These same characteristics, however, are also responsible for the deficiencies in the behavior of the network. For instance, the presence of a distributed form of nonlinearity and the high connectivity of the network make the theoretical analysis of a multilayer perceptron difficult. The learning process must decide which features of the input pattern should be represented by the hidden neurons. It also needs to compute large number of functions, and a choice has to be made between alternative representations of the input pattern.

Chapter 6 Application of PCA & ANNs for Evaluation of Filler Dispersion

6.1 Introduction

The objective of this research is to find a way to correlate the measured data with the dispersion state of the filler to determine the state of the filler dispersion. A model relating the measured variables and the filler dispersion is needed. However, at the current state of the art, there is no established theoretical relation between the measured variables used in the study and the state of filler dispersion, mainly due to the lack of understanding of the complex interaction between ultrasound and filled polymers under flowing condition. Artificial neural network approach provides a way of modeling the relationship between the measured variables and the state of a complex process without fully understanding the process. In order to build an efficient neural network model, the principal component analysis (PCA) was first applied to the multivariate data set collected by multiple sensors in the experiment. The application of PCA allows determining several uncorrelated feature parameters most sensitive to the filler dispersion state. Then a neural network model with these feature parameters as inputs was established.

6.2 Experiment data and the application of PCA

The experiments were carried out using the equipment described in Chapter 3.

6.2.1 Original data

A wide range of data was collected. The experiment data used for modeling and analysis are presented in the appendix. A summary of the variables is given in Table 6.1.

Table 6.1 Summary of process variables

Variables	Description
Q_{feed}	feeding rate of materials (kg/hr)
V_{us}	ultrasound velocity (m/s)
α_{us}	ultrasound attenuation coefficient (dB/cm)
A_{mps}	amperage of electric current to drive the screws of extruder (A)
T_{melt}	melt temperature at the ultrasonic probes location (°C)
P_1, P_2, P_3	pressures read at three pressure probe locations at the die (psi)
P_{us}	pressure at the ultrasonic probes location (psi)
S_{stress}	shear stress (psi)
F_{type}	surface condition of filled mineral particles
V_{RPM}	screw rotation speed (RPM)
T_{prfl}	barrel temperature profile
C_f	filler concentration (wt%)

The controlled parameters of the process were feeding rate Q_{feed} (2 levels: 3.5 and 7.5 kg/h), screw rotation speed V_{RPM} (3 levels: 100, 175, and 300 rotations per minute), barrel temperature profile T_{prfl} (2 profiles: one starting at 185°C at feed throat, with a

gradual increase of 5°C from barrel to barrel up to 225°C at the ninth, with that set-point constant for the remaining barrels and the die; another one with a constant set-point of 200°C imposed for all sections), the type of filler fed into the extruder F_{type} (2 types: one with stearate coating, one without), and the filler concentration C_f (5 levels: 0, 5, 10, 15, and 20 wt%).

The measured parameters were ultrasound speed V_{us} and attenuation coefficient α_{us} in the extruded material, the melt temperature T_{melt} read from the melt thermocouple, the amperage of electric current A_{mps} required to drive the screws of the extruder, the pressures P_1, P_2, and P_3 at three pressure probe locations at the instrumented die, the pressure at the ultrasonic probes location P_{us}, which was an extrapolated value of the pressure readings P_1, P_2, and P_3, and finally the shear stress S_{stress}. The shear stress was determined by:

$$S_{stress} = \frac{H}{2(1+H/W)}\frac{dP}{dx} \tag{6.1}$$

where H is the height and W is width of the rectangular slot channel in the instrumented die, and dP/dx is the pressure gradient in the slot along the melt flow direction. The value of dP/dx was computed based on readings of P_1, P_2, and P_3 and the distances between the pressure probes. The parameter of interest to this study was filler dispersion index D_x.

The surface condition, F_{type}, i.e., with or without stearate coating of fillers, had strong effects on the measurement data [22]. The state of dispersion was examined using SEM. Dispersion index was defined by Eqs. (2.1) and (2.2). The dispersion is best when no agglomerate is detected (D_x = 1). In the case of worst dispersion, all the particles

remain in the form of agglomerates and D_x take the value of 0. In the present study, the critical diameter was set to 7 μm (resolution of the SEM photographs). The size of the examined area by SEM was 1.125mm by 1.125mm. Details on laboratory measurements of filler dispersion can be found in [22].

6.2.2 Application of PCA

In this work, eight directly measured variables, namely, α_{us}, V_{us}, Q_{feed}, T_{melt}, A_{mps}, P_1, P_2, and P_3, plus two derived variables, namely P_{us} and S_{stress} were used. Application of PCA to these 10 variables resulted in 10 principal components (PCs). Table 6.2 presents the correlation coefficient between each original variable and the measured D_x. Table 6.3 shows the correlation coefficient, γ, between each PC and the off-line measured dispersion index D_x.

Table 6.2 Correlation coefficient between original variables and D_x

Variables	α_{us}	V_{us}	P_2	P_{us}	T_{melt}
γ	-0.39	0.009	-0.48	-0.56	-0.29
Variables	A_{mps}	S_{stress}	P_1	P_3	Q_{feed}
γ	-0.65	-0.55	-0.54	-0.51	-0.57

Table 6.3 Correlation coefficient between PCs and D_x

PCs	*1*	*2*	*3*	*4*	*5*
γ	0.55	-0.40	0.17	-0.33	0.07
PCs	*6*	*7*	*8*	*9*	*10*
γ	-0.06	0.20	-0.11	0.01	-0.07

To better visualize the results in Table 6.2 and 6.3, the correlation coefficients are plotted in Fig. 6.1 and 6.2.

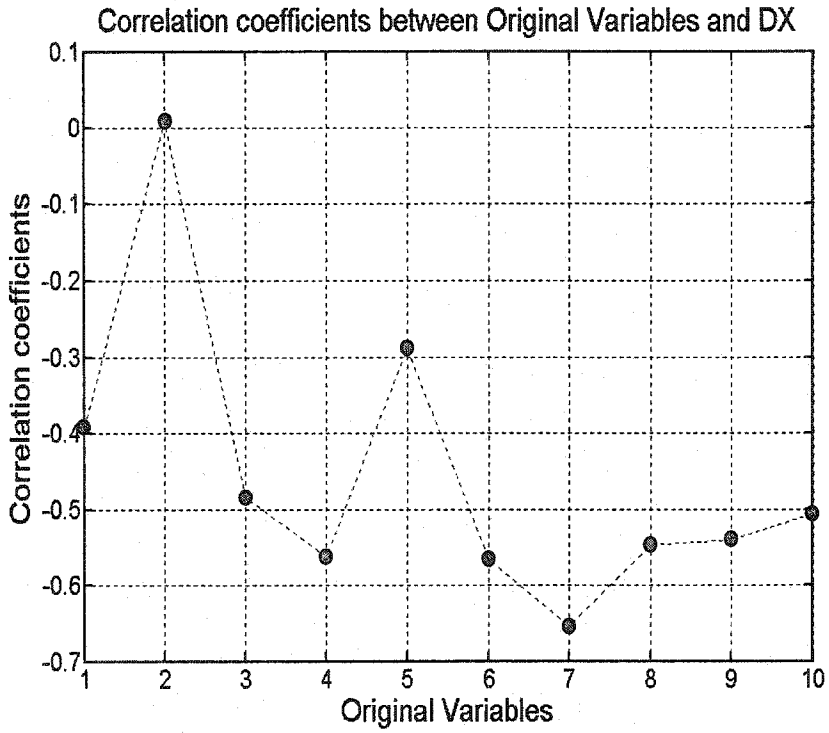

Fig. 6.1 Plot of correlation coefficients between original 10 variables and D_x

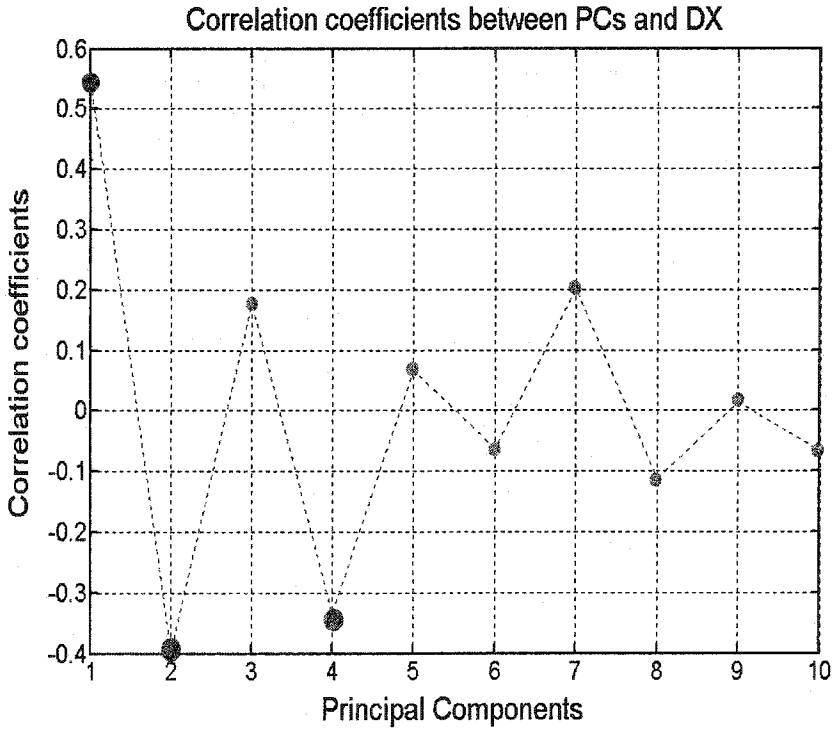

Fig. 6.2 Plot of correlation coefficients between PCs and D_x

6.2.3 Analysis of PCA application

First we look at the correlation coefficients between the input original data and the dispersion index shown in Fig. 6.1. The result indicates that 8 out of the10 variables were correlated with D_x with an absolute correlation coefficient larger than 0.3. Also from Table 6.4, we know that there exist correlations among these variables. For example, all variables except V_{us}, α_{us}, and T_{melt} were significantly correlated with A_{mps}. If choosing all these eight variables, it will increase the complexity of the analysis. Therefore in this situation, it was difficult to decide how many variables should be retained as the input of the system model.

Table 6.4 Correlation coefficients between original variables

	Q_{feed}	V_{us}	α_{us}	A_{mps}	T_{melt}	P_1	P_2	P_3	P_{us}	S_{stress}
Q_{feed}	1	-0.040	0.326	0.870	0.063	0.774	0.800	0.829	0.752	0.732
V_{us}	-0.040	1	-0.541	0.230	-0.685	0.424	0.446	0.430	0.408	0.413
α_{us}	0.326	-0.541	1	0.269	0.277	0.304	0.247	0.238	0.327	0.331
A_{mps}	0.870	0.230	0.269	1	-0.013	0.863	0.877	0.892	0.843	0.830
T_{melt}	0.063	-0.685	0.277	-0.013	1	-0.371	-0.400	-0.357	-0.352	-0.371
P_1	0.774	0.424	0.304	0.863	-0.371	1	0.982	0.975	0.996	0.994
P_2	0.800	0.446	0.247	0.877	-0.400	0.982	1	0.996	0.961	0.957
P_3	0.829	0.430	0.238	0.892	-0.357	0.975	0.996	1	0.951	0.943
P_{us}	0.752	0.408	0.327	0.843	-0.352	0.996	0.961	0.951	1	0.999
S_{stress}	0.732	0.413	0.331	0.830	-0.371	0.994	0.957	0.943	0.999	1

As can be seen in Table 6.2, among the 10 PCs, the 1st, 2nd and 4th PCs are most correlated to D_x, and therefore they were chosen to determine D_x. Also from Fig. 6.2, one can find that only PC1, PC2 and PC4 (plotted with large dot in the figure), are correlated with the dispersion state having absolute correlation coefficients larger than 0.3. This

indicates that these three PCs are most sensitive to the dispersion state of the system. It demonstrates that using PCA has reduced the number of D_x-sensitive variables from 8 to 3 and improved the efficiency of the analysis. We also note that although PC3 has more system information of the original data set than PC4, it is less sensitive to the filler dispersion state than PC4.

6.3 Application of artificial neural network

Artificial neural network was constructed for further analysis based on results from PCA.

6.3.1 Network structure

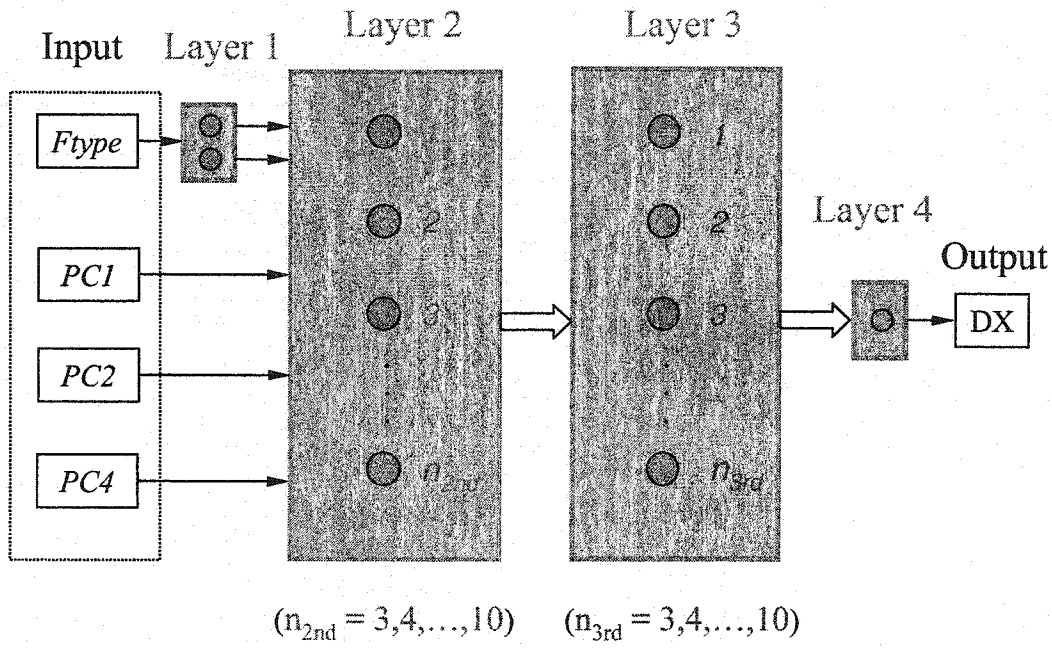

Fig. 6.3 Four-layer feed-forward networks for the estimation of filler dispersion.

A feed-forward network model was constructed for estimating the dispersion index as illustrated in Fig. 6.3. Previous study showed that the surface condition, i.e., with or without stearate coating of fillers, had strong effects on the measurement data. Therefore, the filler type F_{type} was used as an input to the network. The filler type was

coded as 1 for the stearate-coated filler and 2 for the non-coated filler. The effects of the two filler types on the process could not be represented with these two rather arbitrarily chosen simple values. We used a two-neuron layer to represent the effect of the filler type. With this layer, each filler type was characterized by two output numbers. Their values are determined by the network training process. The 1^{st}, 2^{nd}, and 4^{th} PCs determined in Section 6.2 were used as inputs to the second layer. The 2^{nd} and 3^{rd} layers composed of various numbers of neurons in the range from 3 to 10 were investigated. Table 6.5 presents the structures of all neural network models tested in this research. The output of the network was the dispersion index D_x. This network used a hyperbolic tangent sigmoid transfer function for the neurons at the hidden layers, and a linear transfer function for the output layer. Only the first hidden layer had a bias connection.

The weight factors for the input and the connection between two neurons as well as the biases of the first layer were adjusted according to back propagation optimization with early stopping strategy. The minimum output error of the network was achieved. Of the available 93 dispersion indices measurements, 47 were used to form a data set for network training. The others were used to form a validation-testing data set. For each training cycle, half of the data samples were chosen randomly from the validation-testing data pool to form a validation data set, and the remaining data formed the testing data set. A total of 15 pairs of the data sets were used in the training process, resulting in 15 trained networks, determined from trainings for 500 different initial weight and bias conditions. The best result has the largest C_{rr}/E_{max} ratio, where C_{rr} is the correlation coefficient between the measured and estimated D_x given by

Table 6.5 Summary of configurations of neural network models

(Two neurons at Layer 1 and one neuron at Layer 4 for all configurations)

Configurations	Structure of NN models		Configurations	Structure of NN models	
	Neurons of Layer2 (n_{2nd})	Neurons of Layer3 (n_{3rd})		Neurons of Layer2 (n_{2nd})	Neurons of Layer3 (n_{3rd})
1	3	3	33	7	3
2	3	4	34	7	4
3	3	5	35	7	5
4	3	6	36	7	6
5	3	7	37	7	7
6	3	8	38	7	8
7	3	9	39	7	9
8	3	10	40	7	10
9	4	3	41	8	3
10	4	4	42	8	4
11	4	5	43	8	5
12	4	6	44	8	6
13	4	7	45	8	7
14	4	8	46	8	8
15	4	9	47	8	9
16	4	10	48	8	10
17	5	3	49	9	3
18	5	4	50	9	4
19	5	5	51	9	5
20	5	6	52	9	6
21	5	7	53	9	7
22	5	8	54	9	8
23	5	9	55	9	9
24	5	10	56	9	10
25	6	3	57	10	3
26	6	4	58	10	4
27	6	5	59	10	5
28	6	6	60	10	6
29	6	7	61	10	7
30	6	8	62	10	8
31	6	9	63	10	9
32	6	10	64	10	10

$$C_{rr} = \frac{\sum_{i,j}^{N} x_i y_j}{\sqrt{\sum_{i}^{N} x_i^2} \sqrt{\sum_{j}^{N} y_j^2}} \tag{6.2}$$

and E_{max} is the maximum of the errors between all the estimated D_x and the measured ones defined by

$$E_{\max} = \mathrm{MAX}\{|\, x_i - y_i \,|\} \tag{6.3}$$

In Eqs. (6.2) and (6.3), x_i and y_i respectively represent estimated D_x and measured D_x for the ith sample, and N is the total number of samples. Among the 93 estimates of D_x, 47 were from the training data set, 23 from the validation data set, and 23 from the testing data set.

6.3.2 Results and analysis

In order to determine the best network structure to the estimation of D_x, the second and third hidden layers with 3 to 10 neurons were tested. In total, 64 network structures were evaluated for different numbers of neurons at the second and third hidden layers. For each network structure, 15 pairs of validation and testing data sets were used. For each pair of these data sets, the network was trained for 500 different initial weight and bias conditions. Overall, 7500 trained networks were obtained for each network structure. Let E_{max_ave}, E_{max_std}, C_{rr_ave}, and C_{rr_std}, to represent respectively the average and standard deviation of E_{max} and C_{rr}. They were calculated by the 7500 trained networks for each network structure tested. Figs 6.4 to 6.7 and Table 6.6 show the values of E_{max_ave}, E_{max_std}, C_{rr_ave}, and C_{rr_std} corresponding to the numbers of neurons at the 2nd and 3rd hidden layers. In the figures, the result obtained with a second layer of n_{2nd} neurons and a third layer of n_{3nd} neurons is indicated with notation (n_{2nd}, n_{3nd}).

The displayed results are arranged first in ascending order of n_{3nd} and then in ascending order of n_{2nd}. In the series of figures, repetitive patterns are seen clearly when

n_{2nd} passing from one number to the next. From Fig. 6.4, one can see that for a given n_{2nd}, the best performance with least average error is achieved by the networks with an n_{3nd} in the range of 4 to 6. Also one can see that the performance improves when n_{2nd} changes from 3 to 5 and deteriorates gradually with the increase of n_{2nd} from 6 to 10. Among the 64 network structures, the (5,5) and (5,6) structures generate the smallest average errors. In Fig. 6.5, one can see that for a given n_{2nd}, the best performance with largest average correlation coefficient is achieved by the networks with n_{3nd} in the range of 5 to 8. Similar to that shown in Fig. 6.4, the best performance improves when n_{2nd} changes from 3 to 7 and deteriorates when n_{2nd} increases from 8 to 10. The ratio of C_{rr_ave}/E_{max_ave} is displayed in Fig. 6.6. Again, the best performance is achieved with the (5,5) and (6,5) structures. The results presented in Figs. 6.4 to 6.6 confirm that there are optimal sizes for the hidden layers of the network. If the size is too small, the network cannot handle the system complexity; if it is too large, the network may not be efficiently trained with limited quantity of data samples to produce robust predication.

Fig 6.7 shows the changes of the standard deviation of the maximum error E_{max} with the sizes of the 2nd and 3rd hidden layers. As can be seen in the figure, the structure (5,5) generates the least standard deviation, suggesting that this structure can produce most consistent predictions under different training conditions. Based on the above discussion, the network with 5 neurons for the second hidden layer and 5 neurons for the third hidden layer was retained to determine the filler dispersion index.

Figure 6.8 and Table 6.7 show the estimation of the dispersion index based on the average of the estimates generated by 5 best-trained networks. The data are arranged in ascending order of the measured dispersion index. The maximum error is 0.043.

Deviations of the estimated dispersion indices from the measured data are due to insufficient data in network training. In addition, the network cannot reproduce the exact process. The instability of the process, the discrepancy between the readings and the real values of the process and measured variables, and the limited accuracy on dispersion index measurement can also result in deviations of the network estimates.

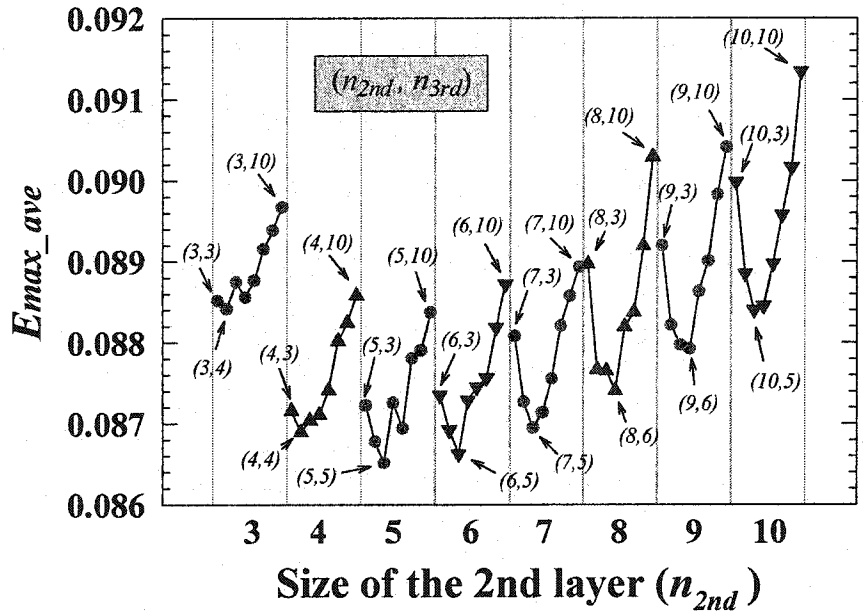

Fig. 6.4 Evolution of the average of maximum errors (E_{max_ave}) with the size of the 2nd hidden layer, n_{2nd}, and the size of the 3rd hidden layers, n_{3rd}

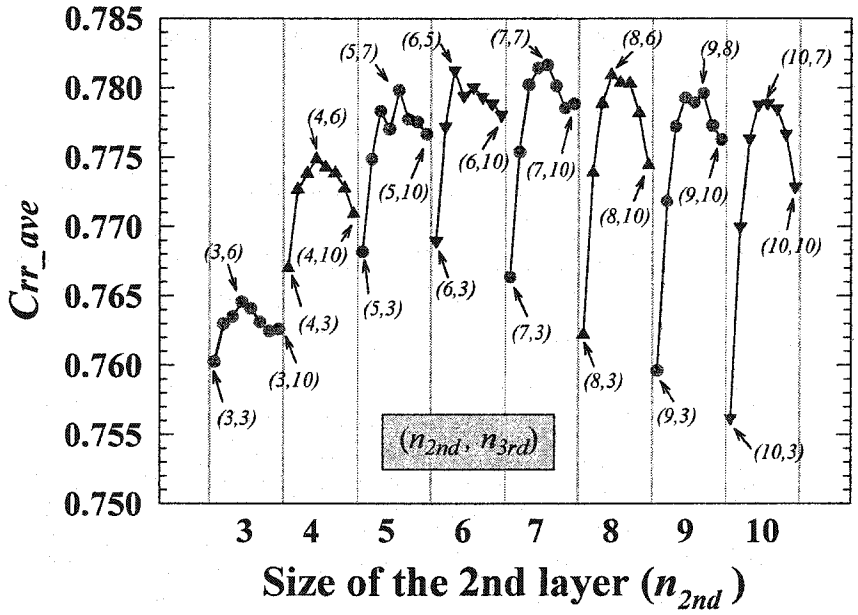

Fig. 6.5 Evolution of the average of the correlation coefficients between the estimated and measured dispersion indices (C_{rr_ave}) with the size of the 2nd hidden layer, n_{2nd}, and the size of the 3rd hidden layers, n_{3rd}

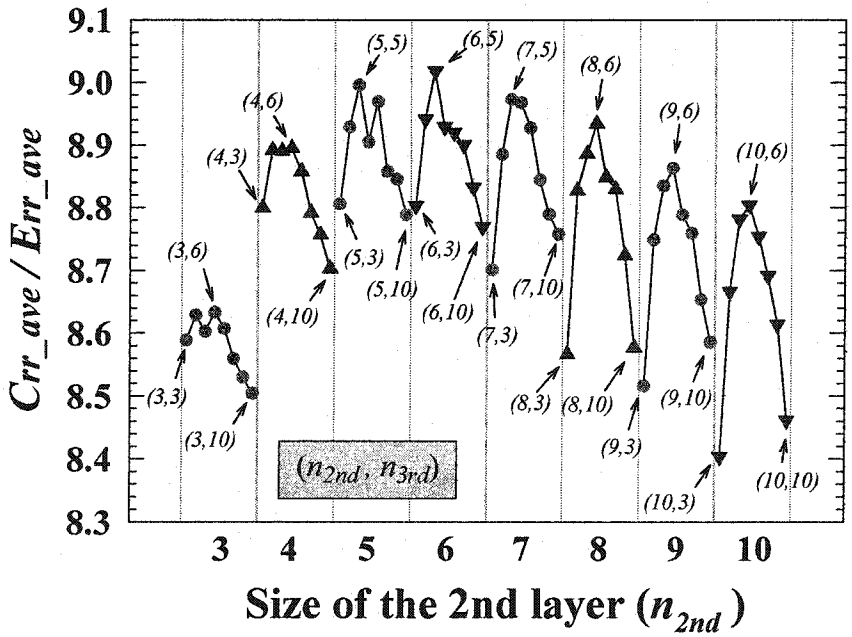

Fig. 6.6 Evolution of C_{rr_ave}/E_{max_ave} with the size of the 2nd layer, n_{2nd}, and the size of the 3rd layers, n_{3rd}

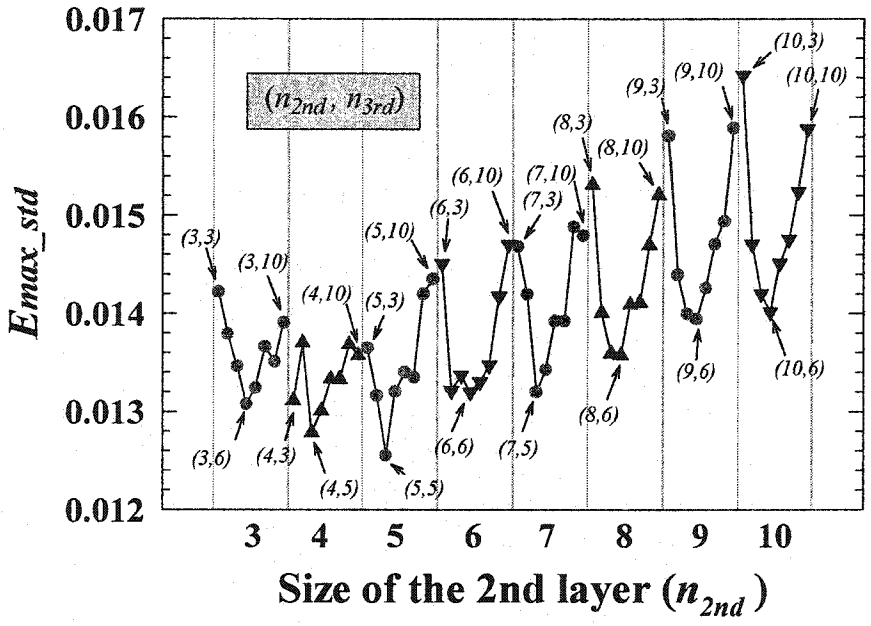

Fig. 6.7 Evolution of the standard deviation of the maximum errors (E_{max_std}) with the size of the 2nd layer, n_{2nd}, and the size of the 3rd layers, n_{3rd}

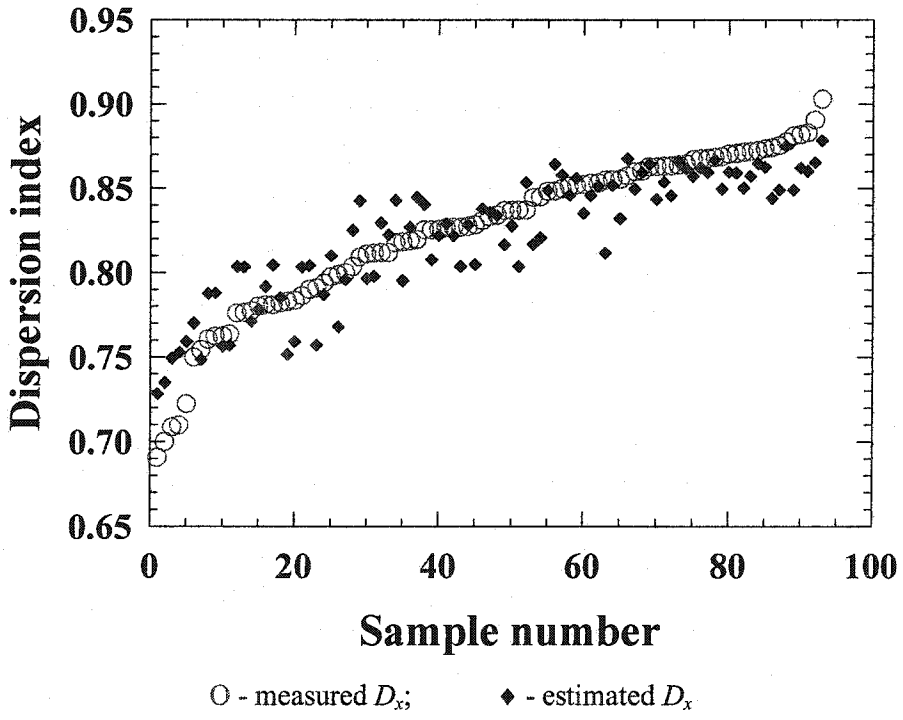

Fig. 6.8 Comparison between the measured (circles) and the estimated dispersion indices produced by the average of the estimates generated by 5 best-trained networks

Table 6.6 Simulation results

Config.	E_{max_ave}	C_{rr_ave}	C_{rr_ave}/E_{rr_ave}	E_{max_std}	Config.	E_{max_ave}	C_{rr_ave}	C_{rr_ave}/E_{rr_ave}	E_{max_std}
1	0.088517	0.76024	8.5886	0.014223	33	0.088082	0.76634	8.7003	0.014684
2	0.088422	0.76298	8.6289	0.013793	34	0.087270	0.77538	8.8849	0.014198
3	0.088750	0.76348	8.6026	0.013464	35	0.086951	0.78020	8.9728	0.013202
4	0.088560	0.76455	8.6331	0.013078	36	0.087141	0.78142	8.9673	0.013428
5	0.088771	0.76408	8.6074	0.013246	37	0.087558	0.78164	8.9271	0.013926
6	0.089157	0.76311	8.5592	0.013663	38	0.088206	0.78013	8.8444	0.013924
7	0.089386	0.76245	8.5299	0.013511	39	0.088577	0.77855	8.7895	0.014883
8	0.089674	0.7626	8.5042	0.013908	40	0.088933	0.77884	8.7576	0.014797
9	0.087159	0.76701	8.8001	0.013113	41	0.08897	0.7622	8.5669	0.015311
10	0.086897	0.77266	8.8917	0.013702	42	0.087667	0.7739	8.8277	0.014008
11	0.087039	0.77383	8.8905	0.012786	43	0.08765	0.77885	8.8859	0.013591
12	0.087112	0.77487	8.8952	0.013009	44	0.087407	0.78092	8.9343	0.013576
13	0.087413	0.77428	8.8578	0.013334	45	0.088195	0.78037	8.8482	0.014092
14	0.088019	0.77385	8.7919	0.013331	46	0.088376	0.78032	8.8295	0.014109
15	0.088246	0.77277	8.7570	0.013685	47	0.089197	0.77817	8.7242	0.014698
16	0.088586	0.77093	8.7026	0.013573	48	0.090291	0.77442	8.5769	0.015207
17	0.087227	0.76816	8.8064	0.013648	49	0.089196	0.75960	8.5160	0.015812
18	0.086781	0.77486	8.9289	0.013162	50	0.088217	0.77181	8.7489	0.014395
19	0.086519	0.77832	8.9959	0.012557	51	0.087972	0.77722	8.8349	0.013996
20	0.087261	0.77702	8.9046	0.013205	52	0.087925	0.77927	8.8629	0.013947
21	0.086941	0.77983	8.9696	0.013261	53	0.088633	0.77894	8.7884	0.014261
22	0.087806	0.77774	8.8575	0.01335	54	0.089006	0.77959	8.7589	0.014707
23	0.087904	0.77755	8.8454	0.01420	55	0.089825	0.77726	8.6530	0.014943
24	0.088375	0.77666	8.7883	0.014355	56	0.090416	0.77625	8.5853	0.015891
25	0.087352	0.76898	8.8033	0.014508	57	0.089982	0.75618	8.4037	0.016423
26	0.086924	0.77722	8.9414	0.024182	58	0.088852	0.76999	8.6660	0.014706
27	0.086626	0.78124	9.0185	0.013368	59	0.088401	0.77631	8.7817	0.014201
28	0.087289	0.77944	8.9294	0.013199	60	0.088456	0.77875	8.8038	0.014015
29	0.087450	0.78004	8.9198	0.013305	61	0.088976	0.77889	8.7539	0.014514
30	0.087562	0.77935	8.9005	0.013472	62	0.089575	0.77850	8.6911	0.014755
31	0.088185	0.77889	8.8325	0.014177	63	0.090165	0.77665	8.6137	0.015239
32	0.088721	0.77804	8.7695	0.014704	64	0.091343	0.77286	8.4612	0.015881

Table 6.7 Estimated and measured dispersion indices

Runs	Estimated	Measured	Runs	Estimated	Measured	Runs	Estimated	Measured
1	0.72829	0.6906	32	0.82954	0.8119	63	0.81175	0.8549
2	0.73491	0.6999	33	0.82249	0.8119	64	0.85185	0.8552
3	0.74928	0.7085	34	0.84302	0.8176	65	0.83213	0.8554
4	0.75270	0.7096	35	0.79532	0.8183	66	0.86759	0.8574
5	0.75917	0.7223	36	0.82696	0.8189	67	0.84961	0.8599
6	0.77012	0.7499	37	0.84462	0.8196	68	0.85939	0.8602
7	0.74855	0.7543	38	0.84056	0.8254	69	0.86433	0.8628
8	0.78763	0.7605	39	0.80763	0.8254	70	0.84358	0.8630
9	0.78802	0.7623	40	0.82203	0.8255	71	0.85368	0.8632
10	0.75631	0.7624	41	0.82904	0.8267	72	0.84575	0.8634
11	0.75704	0.7635	42	0.82202	0.8268	73	0.86607	0.8640
12	0.80368	0.7760	43	0.80383	0.8270	74	0.86179	0.8645
13	0.80332	0.7763	44	0.82838	0.8274	75	0.85721	0.8671
14	0.77127	0.7772	45	0.80493	0.8283	76	0.86225	0.8676
15	0.77825	0.7802	46	0.83781	0.8309	77	0.85947	0.8678
16	0.79186	0.7809	47	0.83599	0.8331	78	0.86709	0.8683
17	0.80453	0.7809	48	0.83392	0.8340	79	0.84957	0.8689
18	0.78525	0.7818	49	0.81666	0.8368	80	0.85986	0.8704
19	0.75159	0.7829	50	0.82772	0.8370	81	0.85919	0.8708
20	0.75912	0.7836	51	0.80385	0.8371	82	0.85026	0.8715
21	0.80338	0.7864	52	0.85335	0.8372	83	0.85741	0.8720
22	0.80421	0.7901	53	0.81674	0.8445	84	0.86481	0.8725
23	0.75711	0.7909	54	0.82090	0.8450	85	0.86260	0.8735
24	0.78706	0.7938	55	0.84882	0.8482	86	0.84420	0.8743
25	0.81002	0.7976	56	0.86430	0.8484	87	0.84904	0.8754
26	0.76792	0.7989	57	0.85792	0.8498	88	0.87535	0.8777
27	0.79600	0.8000	58	0.84600	0.8513	89	0.84911	0.8812
28	0.82528	0.8036	59	0.85608	0.8517	90	0.86218	0.8820
29	0.84255	0.8093	60	0.83530	0.8523	91	0.86030	0.8830
30	0.79663	0.8110	61	0.84575	0.8531	92	0.86525	0.8906
31	0.79783	0.8115	62	0.85140	0.8532	93	0.87851	0.9030

(Note: the data has been sorted according to the ascending order of measured D_x)

6.4 Summary

In this chapter, we built a four-layer feed-forward network model to predict mineral filler dispersion state. The data were collected by a multi-sensor system in the experiment. In order to achieve a precise prediction, principal component analysis (PCA) was first applied to the measured data. 10 measured variables were chosen as the inputs of PCA, resulting in 10 principal components (PCs). It is to be noted that all these PCs are uncorrelated. Through calculating the correlation coefficients between the resulted 10 PCs and the dispersion index, D_x, three PCs most sensitive to the filler dispersion were selected as inputs of neural network models. This allowed us to reduce the redundancy of information and improve the efficiency of the analysis.

Sixty-four four-layer feed-forward network models with one input layer, one output layer, and two hidden layers with number of neurons in each layer ranging from 3 to 10 were constructed and tested to decide the best network structure for estimating D_x. We used 15 pairs of validation-testing data sets for each network structure. For each pair of these data sets, the network was trained for 500 different initial weight and bias conditions. Consequently, 7500 trained networks were obtained for each network structure. The results indicated that there existed repetitive patterns as the number of neurons of the hidden layers changed. And there were optimal sizes for the hidden layers of the network. According to the calculation, the network with 5 neurons for the 2^{nd} hidden layer and 5 neurons for the 3^{rd} hidden layer performed best and was retained for the determination of filler dispersion index. The final simulation results showed that the maximum error between the estimated and the measured filler dispersion index was only

0.043. Considering the limited data size and measurement accuracy, the performance of the network was encouraging. This demonstrated that artificial neural network is a good alternative to physical modeling of the relationship between directly measurable variables and physical properties of a material of interest for process-monitoring purposes.

Chapter 7 Summary and Conclusion

7.1 Summary

Due to varying operational conditions, the polymer extrusion process must be monitored continuously to ensure a reliable and efficient operation. In this work, a multi-sensor system consisting of ultrasonic sensors, thermocouples, pressure sensors and an amperometer was used to monitor the extrusion of a calcium carbonate-filled polypropylene system. A method combining principal component analysis (PCA) and artificial neural networks has been developed and applied to determine the dispersion index of the filler based on the measured data. This hybrid method was developed and applied to the polymer dispersion monitoring. The artificial neural network approach was used since it can establish relationship between the measured data and the state of the filler dispersion without completely understanding the process. In order to build an efficient neural network model, PCA was first applied to the multivariate data set collected by the multiple sensors. The application of PCA had allowed finding out several uncorrelated feature parameters most sensitive to the filler dispersion state. A large number of ANNs with different configurations were tested and the best one was chosen for actual prediction. Taking into consideration the limited samples and data inaccuracy in this research, the performance of the neural network model was very good.

7.2 Contributions

The results of this research suggest that the proposed PCA/ANNs approach, combined with the use of the multi-sensor system, can be used to provide accurate on-line

estimation of filler dispersion during polymer extrusion processes. The application of this method will facilitate the control of polymer extrusion processes and lower production cost. It should be pointed out that this study has been intended to be a part of the work for establishing a hybrid PCA and ANNs approach for effectively extracting process information generated by a multi-sensor system. It is our belief that this approach can be applied to more complex sensing systems. This method provides a way to effectively extract information from the measured data generated by ultrasound, Ramon, dielectric and infrared measurements. Due to the trend of heavier instrumentation and larger data quantity recorded in polymer industry, we anticipate the adoption of this approach by practitioners in the near future.

7.3 Future research

This study was to predict the filler dispersion state from the measured variables and the final result was excellent. In the future, the research in this area can be directed to:

1. Machine-dependent and machine-independent model. In this research, the purpose of applying all the measured variables was to establish the feasibility of using multi-sensor system for monitoring extrusion processes. Among the 10 measured variables, some were machine-dependent, such as temperature profile, amperage of driving motor, and others were machine-independent, such as ultrasonic velocity and attenuation coefficient. The model established in this work is extruder-dependent one. In other words, if the configuration of screws or extruder changes, the model will no longer be valid. Given that it takes much effort to analyze the samples, a machine or extruder independent model will be very much desirable for practical applications. The future

work may aim at building a machine-independent model for polymer extrusion processes. Also a model based only on controlled variables for optimal control can be the future research target.

2. The accuracy of the ANN model developed in this work may be improved by other process matrices as modeling inputs, or by other learning algorithms. The problems associated with artificial neural networks, such as sensitivity to model initialization, training strategy, and number of neurons for a given number of inputs, etc. need further studied.

3. One may give a broad perspective on the value of the combined PCA and ANNs approach by comparing it with other chemometrics techniques, such as partial least square (PLS).

References

1. Anderson, J. A. and Rosenfeld, E., *Neurocomputing: Foundations of Research*. MIT Press, Cambridge, MA, 1988.
2. Anderson, T. W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, New York, 1984.
3. Appello, M., "Characterization of dispersion of carbon black in high density polyethylene using dielectric measurements", *Plastics, Rubber and Composites*, vol. 29, pp.207-211, 2000.
4. Baird, D. G. and Collias, D. I., *Polymer Processing: Principal and Design*, Willey, New York, 1998.
5. Barron, A.R., "Statistical properties of artificial neural networks", *In Proceedings Of the 28th Conference on Decision and Control*, Tampa, Florida, USA, pp.280-285, Dec. 1989.
6. Bertsekas, D.P., *Nonlinear Programming*, Athenas Scientific, Belmont, MA, 1995.
7. Binet, D., *Quantification de la dispersion d'une charge minérale et d'un polymère extrudés par les réseaux de neurones*, Master thesis, University Sherbrooke (Quebec), Canada, 1996.
8. Bories, M., Huneault, M.A., Lafleur, P.G., "Mixing/screw extrusion - effect of twin-screw extruder design and process conditions on ultrafine CaCO3 dispersion into PP", *International Polymer Proceedings*, XIV 3, pp.234-240, 1999.
9. Bridge, B. and Cheng, K.H., "On-line ultrasonic monitoring of the extrusion of CaCO3-filled polypropylene", *Journal of Materials Science Letter*, vol. 6, pp.219-222 1987.
10. Bridge, B. and Cheng, K.H., "On-line ultrasonic monitoring of the extrusion of CaCO3-filled polypropylene", *Journal of Materials Science Letter*, vol. 6, pp.219-222, 1987.
11. Brown, E.C., Collins, T.L.D., Dawson, A.J., Olley, P. and Coates, P.D., "Ultrasound: a virtual instrument approach for monitoring of polymer melt variables", *SPE ANTEC Technique Papers*, vol. 44, pp.335-339, 1998.

12. Bulsari, A. B., *Neural Networks for Chemical Engineers*, Elsevier, 1995.

13. Callister Jr., W. D., *Materials Science and Engineering: An Introduction*, John Willey & Sons, New York, 1994.

14. Damarla, T. R., Ghosal, S., and Karpur, P., "Application of neural networks for classification of ultrasonic signals", *Intelligent Engineering Systems Through Artificial Neural Networks* (Eds Dagli, C.H., Kumara, S.R.T. and Shin, Y.C.) ASME Press, New York, pp.377-382, 1991.

15. Damarla, T. R., Karpur, P. and Bhagat, P.K., "A self-learning neural net for ultrasonic signal analysis", *Ultrasonics*, vol. 30, pp.317-324, 1992.

16. Davis, J., *Statistics and Data Analysis in Geology*, Wiley, New York, 1986.

17. Dobbin, C.J.B. and Baker, W.E., "Analysis of dispersion quality in highly pigmented polymer systems using scanning electron microscopy and image analysis techniques", *50th SPE, ANTEC* 1992, Detroit, Mich, USA, pp.909-912, 1992.

18. Erwin, L. and Dohner, J., "Measurement of mixing on polymer melts by focused ultrasound", *Polymer Engineering Science*, vol. 24, pp.1277-1282, 1984.

19. Ess, J.W. and Hornsby, P.R., "Twin-screw extrusion compounding of mineral filled thermoplastics: dispersive mixing effects", *Plastics and Rubber Processing and Applications*, vol. 8, pp.147-156, 1987.

20. Franca, D. R., Jen, C.-K., Nguyen, K.T. and Gendron, R., "Ultrasonic on-line monitoring of polymer extrusion", *Polymer Engineering and Science*, vol.40, pp.82-94, 2000.

21. Gasteiger, J. and Zupan, J., "Applications of neural networks in chemistry", *CICS Japan Bulletin*, 9, pp.14-23, 1991.

22. Gendron, R. and Binet, D., "State of dispersion: polypropylene filled with calcium carbonate", *Journal of Vinyl & Additive Technology*, vol. 4, pp.54-59, 1998.

23. Goodacre, R., "Characterization and quantification of microbial systems using pyrolysis mass spectrometry: introducing neural networks to analytical pyrolysis", *Microbiology Europe*, vol.2, pp.16-22, 1994.

24. Graham, I.S., Piche, L. and Grant, M., "Experimental evidence for localization of acoustic waves in three dimensions", *Physical Review Letter*, vol. 64, pp.3135-3139, 1990.

25. Guez, Y., Donohue, K.D. and Bilgutay, N.M., "A neural network architecture for ultrasonic nondestructive testing", *IEEE 1991 Ultrasonics Symposium*, pp.777-780, 1991.

26. Hopfield, J.J., "Neural networks and physical systems with emergent collective computational abilities", *In Proceedings of the National Academy of Science*, USA vol. 79, pp.2554-2558, 1982.

27. Hotelling, H., "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, vol. 24, pp.417-441, 1933.

28. Hung, B.-N. and Goldstein, A., "Acoustic parameters of commercial plastics", *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, pp.249-254, 1983.

29. Hunter, J.L. and Derdul, P.R., "Acoustic absorption and dynamic viscosity in a long chain polymer", *Journal Acoustic Society America*, vol.42, pp. 1041-1044, 1967.

30. Hush, D.R. and Horne, B.G., "Progress in supervised neural networks", *IEEE Signal Processing Magazine*, pp.8-39, Jan. 1993.

31. Jackson, J. E., *A User's Guide to Principal Components*, Willey, New York, 1991.

32. Jeffers, J.N.R., "Two case studies in the application of principal component analysis", *Applied Statistics*, vol.16, pp. 225-236, 1967.

33. Jen, C.-K., Franca, D. R., Sun, Z. and Ihara, I., "Clad polymer buffer rods for polymer process monitoring", *Ultrasonics*, vol.39, pp.81-89, 2001.

34. Jen, C.-K., Sun, Z., Kobayashi, M. and Shih, C.-K., "Application of ultrasound in the determination of fundamental extrusion performance: barrel and screw wear measurement", *ANTEC* 2002.

35. Jiang, Q., "Principal component analysis and neural network based face recognition", M.S. thesis. http://people.cs.uchicago.edu/~qingj/ThesisHtml/, 1996.

36. Kendall, M., *Multivariate Analysis*, Macmillan, New York, 1980.

37. Lawrence, S., Giles, C.L., Tsoi, A. C. and Back, A. D., "Face recognition: a convolutional neural network approach," *IEEE Transactions on Neural Networks*, vol.8, pp.98-113, 1997.

38. Lee, Y.G., Manas-Zloczower, I. And Feke, D.L., "Analysis of titanium dioxide agglomerate dispersion in linear low density polyethylene and resulting properties of compounds", *Polymer Engineering and Science*, vol. 35, pp.1037-1045, 1995.

39. Li, L. and Masuda, T., "Effect of dispersion of particles on viscoelasticity of CaCO3-filled polypropylene melts", *Polymer Engineering and Science*, vol. 30, pp. 841-847, 1990.

40. Lippmann, R.P., "An introduction to computing with neural nets", *IEEE ASSP Magazine*, vol. 4, pp.4-22, Apr. 1987.

41. Li, Y., Wang, X., Fu, Y. and Ding, W., "Pattern recognition of properties of alumina-magnesia castable and optimization of its technological parameters", *Anhui Gongye Daxue Xuebao* (*Journal of Anhui University of Technology*) (China), vol. 18, pp. 104-106, 2001.

42. Lopez A.A., Rees, D.E., Socas-Navarro, H. and Lites, B.W., "Pattern recognition techniques and the measurement of solar magnetic fields", *Astronomical data analysis*, San Diego, CA, Aug. 2-3, 2001, Bellingham, WA, Society of Photo-Optical Instrumentation Engineers, pp. 96-106, 2001.

43. McCulloch, W.S. and Pitts, W., "A logical calculus of ideas immanent in nervous activity", *Bulletin of Mathematics Biophysics*, vol. 5, pp.115-133, 1943.

44. Müller, B. and Reinhardt, J., *Neural Networks: An Introduction*, Springer-Verlag, Berlin, 1991.

45. Mutagahywa, B.M. and Hemsley, D.A., "A photometric method for the rapid indication of carbon black dispersion in rubber", *Plastics and Rubber Processing and Application*, vol. 5, pp.219-227, 1985.

46. Nielsen, L. E. and Landel, R. F.: *Mechanical Properties of Polymers and Composites*, Marcel Dekker, New York, 1994.

47. Osswald, T.A. and Menges, G., *Materials Science of Polymers for Engineers*, Hanser, New York, 1995.

48. Pearson, K., "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine*, vol.6, pp.559-572, 1901.

49. Piche, L. and Hamel, A., "Ultrasonic evaluation of filled polymers: I. Techniques and models for the elastic moduli of a resin filled with iron spherical inclusions", *Polymer Composites*, vol. 7, pp.356-362, 1986.

50. Piche, L. and Hamel, A., "Ultrasonic evaluation of filled polymers: II. Elastic moduli of a resin filled with iron inclusions of different aspect ratios", *Polymer Composites*, vol. 8, pp.22-28, 1987.

51. Piche, L., "Ultrasonic probe for on-line process monitoring", *SPE ANTEC Technique Papers*, vol. 44, pp.3617, 1998.

52. Piche, L., Hamel, A., Gendron, R., Dumoulin, M.M. and Tatibouët, J., U.S. Patent 5,433,111, 1995.

53. Polanski, J., Gasteiger, J., Wagener, M. and Sadowski, J., "The comparison of molecular surfaces by neural networks and its application to quantitative structure activity studies", *Quantitative Structure-Activity relationships*, vol.17, pp.27-36, 1998.

54. Progelhof, R.C., and Throne, J.L., *Polymer Engineering Principles: Properties, Processes, and Tests for Design*, Hanser Publishers, New York, 1993.

55. Quantin, J.-C., Averous, L. and Crespy, A., "Determination of the microtexture of reinforced thermoplastics by image analysis", *Composites Science and Technology*, vol. 58, pp.377-387, 1998.

56. Ramos, L.S., Beebe, K.R., Carey, W.P., Sanchez, E., Erickson, B.C., Wilson, B. E., Wangen, L.E. and Kowalski, B.R., "Chemometrics", *Analytical Chemistry*, vol. 58, pp.294-315, 1986.

57. Ripley, B.D., "Neural networks and related methods for classification", *Journal of the Royal Statistical Society*, B, vol. 3, pp.409-456, 1994.

58. Rosato, D. V., *Designing with Reinforced Composites*, Hanser, Munich, 1997.

59. Rumelhart, D.E., Hinton, G.E. and Williams, R.J., "Learning representations of back-propagation errors," *Nature* (London), vol.323, pp.533-536, 1986a.

60. Rumelhart, D.E., McClelland, J.L., and the PDP Research Group, "Parallel distributed processing. Experiments in the microstructure of cognition", Vols. I & II, MIT Press, Cambridge, MA, 1986.

61. Sahnoune, A. and Piche, L., "Ultrasonic study of anharmonicity in amorphous polymers: temperature, pressure and molecular weight effects", *Journal of Non-Crystalline Solids*, vol. 235-237, pp.664-669, 1998.

62. Sahnoune, A., Massines, F. and Piche, L., "Ultrasonic measurement of relaxation behavior in polystyrene", *Journal of Polymer Science*, vol. 34, pp.341-348, 1996.

63. Sarle, W.S., ed. (1997), Neural Network FAQ, part 1 of 7: Introduction, periodic posting to the Usenet newsgroup comp.ai.neural-nets, URL: ftp://ftp.sas.com/pub/neural/FAQ.html

64. Serra-Ricart, M., Calbet, X., Garrido, L. and Gaitan, V., "Multidimensional statistical analysis using artificial neural networks – astronomical applications", *Astronomical Journal*, vol. 106, pp. 1685-1695, 1993.

65. Shen, J., Edwards, R., Thomas, C.L. and Bur, A.T., "Ultrasonic melt temperature measurement during extrusion", *SPE ANTEC Technique Papers*, vol. 44, pp.2076-2079, 1998.

66. Singh, H.P., Gulati, R.K. and Gupta, R., "Stellar spectral classification using principal component analysis and artificial neural networks", *Royal Astronomical Society*, Monthly Notices, vol. 295, pp. 312-318, 1998.

67. Suetsugu, Y., "State of dispersion – mechanical properties correlation in small particle filled polymer composites", *International Polymer Proceedings*, vol. 3, pp.184-190, 1990.

68. Sun, Z., Jen, C.-K. and Shih, C.-K., "Application of ultrasound in the determination of fundamental extrusion performance: monitoring of melting and mixing processes", *Polymer Processing Society*, 18th Annual Meeting, Guimarães, Portugal, June 17-20, 2002.

69. Sun, Z., Jen, C.-K., Derouri, A., Tatibouët, V, "Application of ultrasound in the determination of fundamental injection molding performance: in-line monitoring at nozzle", *Polymer Processing Society*, 18th Annual Meeting, Guimarães, Portugal, June 17-20, 2002.

70. Sun, Z., Jen, C.-K., Franca, D. R. and Liaw, J.-W., "Polymer extrusion and cure monitoring using clad polymer buffer rods", *IEEE Ultrasonics Symposium*, San Juan, Puerto Rico, Oct. 22-25,2000.

71. Sun, Z., Tatibouët, J. and Jen, C.-K., "Application of ultrasound in the determination of fundamental extrusion performance: monitoring of single-screw compounding extrusion", *Polymer Processing Society*, 18th Annual Meeting, Guimarães, Portugal, June 17-20, 2002.

72. Tatibouët, J. and Huneault, M.A., "In-line ultrasonic monitoring of filler dispersion during extrusion", *International Polymer Proceedings*, XVII 1, pp.49-52, 2002.

73. Tatibouët, J. and Piche, L., "Ultrasonic investigation of semicrystalline polymers: study of poly (ethylene terephthalate)", *Polymer*, vol. 32, pp.3147-3151, 1991.

74. Verdier, C., Longin, P.-Y. and Piau, M., "Dynamic shear and compressional behavior of polydimethylsiloxanes", Ultrasonic and Low Frequency Characterization, *Rheologica Acta*, vol. 37, pp.234-244, 1998.

75. Wah, C.A., Choong, L.Y. and Neon, G.S., "Effects of titanate coupling agent on rheological behavior, dispersion characteristics and mechanical properties of talc filled polypropylene", *European Polymer Journal*, vol. 36, pp. 789-801, 2000.

76. Wang, H., Jen, C.-K., Nguyen, K.T. and Viens, M., "On-line ultrasonic monitoring of the injection molding process", *Polymer Engineering Science*, vol.37, pp.363-376, 1997.

77. Werbs, P., *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD. Thesis, Applied Math., Harvard University, Nov. 1974.

78. Wise, B.M. and Kowalski, N.B., "Process Chemometrics" in *Process Analytical Chemistry, Blackie Academic & Professional*, London, pp.259-312, 1995.

79. Zupan, J. and Gasteiger, J., "Neural networks: a new method for solving chemical problems or just a passing phase?" *Analytica Chimica Acta*, vol.248, pp.1-30, 1991.

80. Zupan, J. and Gasteiger, J., *Neural Networks in Chemistry and Drug Design*, Second Edition, Wiley-VCH, Weinheim, 1999.

81. Zupan, J., Novic, M. and Gasteiger, J., "Neural networks with counter-propagation learning strategy used for modeling", *Chemometrics and Intelligent Laboratory System*, vol. 27, pp.175-187, 1995.

Appendix: Experimental data (Source: Ref. [7])

Sample index	F_{type}	C_f(%)	T_{prfl}	V_{RPM} (RPM)	Q_{Feed} (kg/h)	V_{us} (m/s)	α_{us} (dB/cm)	A_{mps} (A)	T_{melt} (°C)	P_1 (psi)	P_2 (psi)	P_3 (psi)	P_{us} (psi)	S_{stress} (psi)	D_x
1	Non_ST	5	1	100	8.8235	886.71	17.096	26.48	231.36	400.56	265.98	139.76	541.890	9098.215	0.6906
2	Non_ST	5	1	300	8.3333	845.26	15.636	18.679	233.25	334.61	223.31	118.97	451.488	7522.670	0.7085
3	Non_ST	0	2	300	9.3750	863.12	18.901	20.343	212.06	415.40	278.04	145.59	559.657	9412.222	0.7623
4	Non_ST	5	1	175	8.8235	865.96	15.763	22.228	232.05	380.38	252.44	132.85	514.724	8634.819	0.7635
5	Non_ST	0	2	100	7.8947	942.02	6.5622	24.082	204.73	418.08	281.77	147.68	561.223	9432.817	0.7763
6	Non_ST	5	1	100	7.8947	899.16	7.6280	23.113	231.07	358.17	238.08	126.78	484.281	8071.783	0.7802
7	Non_ST	0	2	300	8.8235	872.49	16.238	19.487	210.27	393.96	264.06	139.12	530.359	8889.796	0.7864
8	Non_ST	0	2	100	4.3750	921.80	14.964	18.758	205.66	373.86	251.68	130.72	502.164	8481.950	0.7976
9	Non_ST	5	1	175	4.3750	851.05	17.808	16.750	230.75	297.78	196.81	103.05	403.814	6793.276	0.8000
10	Non_ST	0	2	100	3.6842	945.68	6.3547	15.718	205.37	327.29	220.66	115.90	439.261	7374.442	0.8093
11	Non_ST	5	1	300	4.3750	825.03	20.586	15.603	232.30	281.53	187.47	98.826	380.318	6373.682	0.8115
12	Non_ST	0	2	100	3.8889	937.15	8.8756	16.579	205.42	339.49	228.81	119.93	455.723	7659.586	0.8196
13	Non_ST	0	2	175	4.3750	896.66	15.265	17.557	207.26	359.44	241.51	125.19	483.295	8171.859	0.8254
14	Non_ST	5	1	100	4.1176	881.97	15.125	17.353	230.57	296.23	195.84	103.54	401.653	6721.917	0.8270
15	Non_ST	0	2	300	4.3750	859.95	18.290	15.463	209.14	315.66	210.80	110.75	425.787	7148.328	0.8283
16	Non_ST	0	2	175	4.1176	903.88	12.901	16.677	206.75	341.34	229.94	120.22	458.319	7713.685	0.8331
17	Non_ST	5	1	300	3.6842	851.39	8.7336	13.323	230.60	223.95	149.72	80.951	301.896	4988.427	0.8498
18	Non_ST	5	1	175	3.6842	879.97	8.1665	13.912	230.65	257.65	172.08	92.540	347.510	5759.737	0.8517
19	Non_ST	0	2	300	3.8889	877.86	11.660	13.988	204.82	278.99	187.20	99.284	375.387	6269.110	0.8599
20	Non_ST	0	2	175	3.6842	919.95	7.4962	14.768	206.2	306.34	206.86	108.73	410.805	6893.497	0.8632
21	Non_ST	5	1	300	3.8889	841.28	13.075	14.479	231.31	243.75	162.80	87.438	328.757	5452.982	0.8754
22	ST	5	1	100	8.3333	890.77	10.826	23.155	230.21	346.13	242.88	132.26	454.563	7461.029	0.7223
23	ST	5	1	100	8.8235	884.17	14.888	23.903	230.41	360.42	253.03	136.92	473.200	7796.792	0.7543
24	ST	5	1	100	7.8947	897.08	6.5878	22.285	229.97	331.07	232.51	127.23	434.562	7110.908	0.7809
25	ST	5	1	100	9.3750	877.28	18.053	24.600	231.21	373.02	261.14	140.26	490.503	8119.646	0.7829
26	ST	5	1	175	9.3750	855.18	15.817	21.408	231.22	340.41	242.40	131.17	443.337	7299.312	0.7909
27	ST	0	2	100	8.8235	921.48	11.479	22.931	205.28	405.97	287.55	151.39	530.330	8880.948	0.8119
28	ST	5	1	175	8.3333	869.49	10.566	20.653	229.54	298.45	226.17	124.25	374.342	6076.790	0.8183

29	ST	0	2	300	7.8947	882.60	10.336	19.854	204.76	331.04	233.82	123.35	433.126	7245.213	0.8267
30	ST	5	1	300	8.8235	833.49	15.774	18.542	233.47	323.15	226.72	122.91	424.409	6985.306	0.8368
31	ST	0	2	100	9.3750	914.18	14.632	24.007	205.62	421.93	297.57	156.76	552.521	9250.359	0.8371
32	ST	5	1	300	9.3750	826.71	17.115	18.576	233.42	333.90	233.99	126.23	438.826	7244.633	0.8445
33	ST	5	1	175	4.3750	850.60	13.121	16.267	230.76	253.88	185.09	95.293	326.132	5532.318	0.8482
34	ST	5	1	100	4.3750	871.25	13.440	15.357	229.24	268.46	196.80	105.41	343.710	5687.943	0.8531
35	ST	5	1	300	8.3333	842.03	13.642	18.933	232.97	314.68	221.15	120.41	412.893	6777.032	0.8549
36	ST	0	2	100	8.3333	929.95	8.0875	22.098	204.88	391.33	277.66	146.30	510.707	8547.942	0.8554
37	ST	0	2	300	3.8889	873.89	8.8515	13.923	205.41	259.65	185.51	97.035	337.517	5672.860	0.8628
38	ST	5	1	175	3.6842	875.60	6.8962	14.714	230.30	225.75	166.35	86.566	288.132	4855.427	0.8640
39	ST	5	1	175	4.1176	858.51	10.984	15.519	230.68	242.81	177.68	91.770	311.193	5268.930	0.8671
40	ST	0	2	175	4.1176	901.31	10.083	15.120	205.01	298.22	211.59	110.17	389.180	6560.055	0.8678
41	ST	0	2	300	8.8235	864.57	13.880	20.396	208.34	358.97	252.93	133.11	470.328	7879.173	0.8689
42	ST	0	2	300	4.3750	855.23	13.196	14.679	206.29	281.96	200.58	104.27	367.439	6198.911	0.8708
43	ST	5	1	100	4.1176	880.27	11.199	14.704	228.93	258.44	190.32	102.62	329.961	5435.868	0.8720
44	ST	0	2	175	3.6842	920.38	6.6897	14.338	204.31	280.69	199.48	104.01	365.988	6163.667	0.8735
45	ST	0	2	175	8.3333	908.79	8.3866	21.110	206.49	373.13	261.92	137.37	489.908	8224.508	0.8812
46	ST	0	2	100	3.6842	942.93	5.5136	15.395	202.93	288.69	208.96	110.18	372.428	6227.629	0.8830
47	ST	5	1	300	3.8889	835.00	9.4031	14.917	231.23	210.63	155.17	81.972	268.871	4488.192	0.9030
48	Non_ST	5	1	100	8.3333	892.70	12.486	24.585	231.15	379.31	252.01	133.44	512.998	8577.094	0.6999
49	Non_ST	5	1	300	9.3750	827.32	21.721	20.039	233.90	370.62	246.27	129.01	501.192	8428.576	0.7499
50	Non_ST	5	1	300	7.8947	854.27	10.817	17.419	232.20	309.80	207.10	110.75	417.651	6943.825	0.7605
51	Non_ST	5	1	175	9.3750	857.94	19.462	23.264	232.54	402.78	266.49	138.74	545.907	9211.054	0.7624
52	Non_ST	0	2	100	8.8235	928.54	14.553	27.405	206.26	457.32	308.17	160.81	613.945	10343.654	0.7772
53	Non_ST	0	2	100	8.3333	934.65	10.406	25.419	205.34	434.95	293.14	153.38	583.878	9822.533	0.7809
54	Non_ST	0	2	175	9.3750	892.57	17.649	24.538	206.00	449.72	302.61	156.24	604.213	10238.066	0.7818
55	Non_ST	5	1	175	8.3333	873.26	12.158	21.319	231.65	359.32	238.78	126.72	485.905	8114.424	0.7836
56	Non_ST	0	2	175	8.8235	900.75	13.946	23.306	204.88	426.68	287.80	149.42	572.536	9672.129	0.7901
57	Non_ST	5	1	100	3.7500	876.06	17.906	18.611	230.98	318.80	210.33	110.44	432.705	7268.709	0.7938
58	Non_ST	5	1	300	8.8235	836.76	18.138	18.952	233.16	348.49	232.21	122.81	470.604	7873.081	0.7989
59	Non_ST	0	2	300	7.8947	891.48	10.866	17.645	208.02	354.45	237.68	126.03	477.073	7968.661	0.8036
60	Non_ST	5	1	175	7.8947	881.16	8.0002	20.179	231.25	339.69	225.84	120.49	459.247	7646.533	0.811
61	Non_ST	0	2	300	3.5526	887.26	8.2033	13.243	206.43	256.85	172.44	92.102	345.494	5747.264	0.8176

62	Non_ST	5	1	175	3.6765	857.84	14.832	15.971	229.76	277.37	183.91	97.299	375.503	6281.583	0.8189
63	Non_ST	5	1	100	3.5526	897.02	7.3819	15.302	230.31	270.18	179.50	96.315	365.408	6065.332	0.8254
64	Non_ST	0	2	175	7.8947	922.32	7.1746	20.160	206.24	397.13	266.04	140.84	534.798	8940.704	0.8255
65	Non_ST	5	1	100	3.6111	889.66	11.388	16.213	230.52	282.39	187.45	99.666	382.102	6374.552	0.8268
66	Non_ST	5	1	300	3.6765	833.03	17.094	14.952	231.73	259.35	172.95	92.203	350.077	5830.806	0.8340
67	Non_ST	0	2	175	8.3333	908.94	10.593	22.369	203.36	405.26	273.55	142.56	543.572	9164.207	0.8370
68	Non_ST	0	2	300	8.3333	881.56	13.980	18.821	209.24	375.28	251.62	133.04	505.137	8450.477	0.8450
69	Non_ST	0	2	100	3.6765	928.27	12.113	17.763	205.58	356.28	240.05	125.35	478.334	8056.119	0.8523
70	Non_ST	0	2	175	3.6111	911.30	10.061	15.566	206.43	319.54	215.68	113.14	428.601	7200.107	0.8552
71	Non_ST	0	2	300	3.6765	868.29	14.433	14.632	206.30	290.26	194.20	102.17	391.138	6561.651	0.8634
72	Non_ST	5	1	225	3.5526	866.41	8.4376	13.706	230.99	252.93	169.04	90.955	341.027	5650.524	0.8676
73	ST	5	1	175	8.8235	862.27	13.474	21.237	230.16	323.33	235.64	128.46	415.402	6798.062	0.7096
74	ST	5	1	175	7.8947	877.28	7.3736	20.332	228.55	292.48	217.55	120.15	371.180	6011.813	0.7760
75	ST	0	2	100	7.8947	939.46	5.3675	21.520	204.65	383.92	272.66	143.75	500.757	8378.248	0.8119
76	ST	0	2	175	9.3750	890.99	13.766	22.415	207.11	407.36	284.90	148.90	535.972	9016.414	0.8274
77	ST	0	2	175	8.8235	900.00	10.766	21.696	206.73	390.85	273.89	143.52	513.666	8628.003	0.8309
78	ST	0	2	100	3.7500	912.72	12.233	16.285	204.22	310.88	223.91	117.68	402.204	6739.612	0.8372
79	ST	5	1	100	3.5526	898.66	6.1621	13.973	228.68	242.93	179.86	97.452	309.162	5074.869	0.8484
80	ST	0	2	300	8.3333	872.70	11.972	19.865	207.06	345.59	243.88	128.70	452.402	7566.182	0.8513
81	ST	0	2	175	3.7500	892.35	12.073	15.846	205.32	308.92	218.78	113.53	403.596	6816.482	0.8532
82	ST	5	1	300	3.7500	818.25	13.786	15.520	231.91	227.77	167.27	86.498	291.308	4928.236	0.8574
83	ST	0	2	300	3.5526	883.42	6.8909	13.550	204.62	252.44	180.85	95.182	327.611	5485.761	0.8602
84	ST	0	2	300	9.3750	856.82	16.362	20.571	209.54	373.72	262.73	137.74	490.271	8232.195	0.8630
85	ST	0	2	100	3.6765	922.37	10.296	15.792	203.99	301.18	217.44	114.54	389.122	6511.177	0.8645
86	ST	5	1	300	3.5526	845.26	6.9809	14.591	230.67	196.42	146.96	77.377	248.348	4152.719	0.8683
87	ST	0	2	300	3.6765	864.14	10.964	14.387	205.88	273.90	195.16	101.75	356.590	6005.286	0.8704
88	ST	0	2	175	7.8947	918.18	6.2375	20.795	206.22	366.36	257.27	135.05	480.931	8069.172	0.8715
89	ST	0	2	100	3.6111	933.15	7.5239	15.569	203.55	291.67	210.93	111.12	376.474	6298.697	0.8725
90	ST	5	1	300	7.8947	848.90	10.998	19.375	232.04	291.04	205.86	113.06	380.477	6208.629	0.8743
91	ST	5	1	300	3.6765	826.38	12.065	15.370	231.50	218.52	161.24	84.083	278.690	4689.939	0.8777
92	ST	5	1	100	3.6111	888.97	8.6863	14.409	228.67	251.02	185.39	100.29	319.953	5258.197	0.8820
93	ST	0	2	175	3.6111	910.54	8.5258	14.737	204.59	288.38	204.81	106.69	376.141	6338.002	0.8906