Chair of Process Technology and Industrial Environmental Protection

Doctoral Thesis

Empirical modeling for the optimized operation and real-time control of coarse shredders for mixed commercial waste

Dipl.-Ing. Karim Khodier, BSc

March 2021

AFFIDAVIT

I declare on oath that I wrote this thesis independently, did not use other than the specified sources and aids, and did not otherwise use any unauthorized aids.

I declare that I have read, understood, and complied with the guidelines of the senate of the Montanuniversität Leoben for "Good Scientific Practice".

Furthermore, I declare that the electronic and printed version of the submitted thesis are identical, both, formally and with regard to content.

Date 17.03.2021

Signature Author
Karim Khodier

DANKSAGUNG

„Wo Gemeinschaft herrscht, da herrscht auch Erfolg"

Diese Worte des römischen Autors Pubilius Syrus beschreiben sehr gut den Geist der vergangenen Jahre, in welchen diese Dissertation entstanden ist. Denn auch wenn sie dazu dient meine persönliche wissenschaftliche Leistung widerzugeben, so hätte sie nicht ohne die Mithilfe vieler bemerkenswerter Menschen entstehen können.

Zuallererst möche ich mich bei meinem Doktorvater, Univ.-Prof. Dipl.-Ing. Dr.-Ing. Markus Lehner, bedanken. Danke dafür, dass Sie mich schon seit dem Bachelorstudium auf vielfältige Art und Weise gefördert haben und Sie mir als Wissenschaftler, als Chef und als Mensch ein Mentor und Vorbild sind.

Mein Dank gilt auch dem zweiten Betreuer meiner Dissertation, O.Univ.-Prof. Dipl.-Ing. Dr.techn. Paul O'Leary. Mit deinen hohen wissenschaftlichen Ansprüchen und deinem mathematisch-datenwissenschaftlichen Wissen hast du wesentlich das methodische Vorgehen meiner Forschungstätigkeit inspiriert.

Ein großes Dankeschön außerdem an Dipl.-Ing. Dr.mont. Renato Sarc, dem Leiter des Forschungsprogramms „Recycling and Recovery of Waste 4.0" (ReWaste4.0), der dieses auch federführend beantragt hat. Danke, dass du den finanziellen und inhaltlichen Rahmen für diese und vier weitere Dissertationen geschaffen hast, uns als Team durch diese Forschungsjahre geführt und uns stets den Rücken gestärkt hast. Außerdem danke für alles was ich von dir über die Interessen verschiedener Steakholder und deren Zusammenführen zu einem gemeinsamen Erfolgsweg lernen durfte.

Ebenso möchte ich mich bei meinen Mit-Doktorandinnen und Mit-Doktoranden Dipl.-Ing. Alexander Curtis, Dipl.-Ing. Selina Möllnitz, Sandra Viczek, MSc MSc und Dipl.-Ing. Thomas Weißenbach bedanken. Danke für den regen Austausch, eure tatkräftige Unterstützung bei den Versuchen und für das tolle Miteinander in den letzten Jahren. Und Danke an meine zahlreichen studentischen Mitarbeiterinnen und Mitarbeiter der vergangenen Jahre – insbesondere seien hier Dipl.-Ing. Lisa Kandlbauer, Jessica Steiner und Maximilian Enengel erwähnt. Erst euer aller unglaublicher Einsatz und eure Eigeninitiative haben die Versuche für diese Dissertation möglich gemacht. Und ohne die gute Stimmung und das lustige und harmonische Miteinander in diesen intensiven Phasen hätten sie nicht annähernd so viel Spaß gemacht.

Danke auch an die Sekretärinnen des Lehrstuhls für Verfahrenstechnik des industriellen Umweltschutzes, Gabriella-Nicola Read und Christa Waltritsch, für eure Hilfsbereitschaft und

ABSTRACT

Empirical modeling for the optimized operation and real-time control of coarse shredders for mixed commercial waste

The European Union strives a transition towards a circular economy, hence requiring effective and efficient treatment processes for the recycling and thermal exploitation of produced waste. Mixed commercial waste presents a significant share of the produced waste, with around 1.1 million metric tons alone in Austria in 2016. It is usually first treated mechanically, with coarse shredders performing the first processing step of comminution and liberation while also functioning as the primary dosing devices. Despite their importance for waste processing, the knowledge about the process behavior of coarse shredders and its dependence on their parametrization is limited. But understanding the influence of different configurable factors is essential for their optimized operation and enhanced processing, involving smart real-time control, in the context of the fourth industrial revolution. Physical numerical models are hardly applicable for understanding these machines' behavior in real-scale operation. They require too detailed information about the always-changing condition of the heterogeneous input material: mixed commercial waste. Consequently, empirical modeling approaches are followed in this thesis. As a necessary precondition for successful experimentation, a procedure for material sampling was first established, based on the theory of sampling. The induced general estimation error was then determined based on a replication experiment. Subsequently, a 32 runs coarse-shredding experiment with mixed commercial waste was conducted, based on a fully randomized, D-optimal experimental design. It aimed at investigating the influences of the radial gap width, the shaft rotation speed, and the cutting tool geometry on a shredder's throughput behavior, its energy demand, and the particle size distribution it produces. Significant models were successfully derived, applying (multivariate) multiple linear regression. For doing so, the particle size distribution was described as isometric log-ratio-transformed mass shares of three particle size classes (>80 mm, 30–80 mm, 0–30 mm). The models show significant effects of all three factors on throughput behavior and energy demand. But only the cutting tool geometry significantly influenced the shares of the particle size classes. Based on the models, conclusions on the optimized operation of coarse shredders were drawn, which contradict common operation settings. Finally, investigations on the sensor-based real-time measurement of particle size distributions, as defined by drum screening, were conducted: partial least squares regression models, based on geometric descriptors, obtained from two-dimensional RGB images of the particles, turned out to be a promising approach.

KURZFASSUNG

Empirische Modellierung für den optimierten Betrieb und die Echtzeit-Regelung von Vorzerkleinerern für gemischte Gewerbeabfälle

Die Europäische Union strebt eine Kreislaufwirtschaft an. Dafür sind effektive und effiziente Behandlungsprozesse für das Recycling und die thermische Nutzung von Abfällen nötig, von welchen gemischte Gewerbeabfälle einen relevanten Anteil darstellen. Sie werden üblicherweise zunächst mechanisch behandelt. Vorzerkleinerer übernehmen dabei den ersten Prozessschritt – Zerkleinerung und Aufschluss – und fungieren zugleich als primäre Dosiereinrichtung. Trotz ihrer Bedeutung für die Abfallbehandlung ist bisher nur wenig über den Einfluss der Vorzerkleinerer-Parametrierung auf deren Prozessverhalten bekannt. Diesen Einfluss zu kennen ist aber essenziell für die Verbesserung der mechanischen Abfallbehandlung und für intelligente Industrie 4.0-Echtzeitregelungen. Physikalische, numerische Modelle sind kaum geeignet, um das Verhalten dieser Maschinen im Industriemaßstab zu verstehen, da sie zu detaillierte Informationen über das sich ständig ändernde Inputmaterial, den Abfall, erfordern. Daher werden in dieser Arbeit empirische Modellierungsansätze verfolgt. Vorbereitend wurden zunächst ein Probenahme-Prozedere, basierend auf der Theorie der Probenahme erarbeitet und verbleibende Probenahmeeinflüsse experimentell quantifiziert. In weiterer Folge wurde ein Vorzerkleinerungsexperiment mit gemischten Gewerbeabfällen, mit 32 Durchläufen, basierend auf einem vollständig randomisierten D-optimalen Versuchsplan durchgeführt. Damit wurden die Einflüsse des radialen Schnittspalts, der Rotordrehzahl und der Zerkleinerungsgeometrie auf das Durchsatzverhalten, den Energiebedarf und auf die erzeugte Korngrößenverteilung untersucht. Unter Anwendung (multivariater) multipler linearer Regression konnten aus den Daten signifikante Modelle abgeleitet werden. Dafür wurde die Korngrößenverteilung durch isometrische Logarithmenverhältnisse der Massenanteile dreier Kornklassen (>80 mm, 30–80 mm, 0–30 mm) beschrieben. Die Modelle zeigen signifikante Einflüsse aller drei untersuchter Faktoren auf das Durchsatzverhalten und den Energiebedarf, wohingegen die Korngrößenverteilung nur von der Zerkleinerungsgeometrie signifikant verändert wurde. Basierend auf den Modellen konnten Schlussfolgerungen bezüglich des optimierten Betriebs von Vorzerkleinerern gezogen werden, welche der gängigen Parametrierung widersprechen. Abschließend wurden Untersuchungen zur sensorgestützten Echtzeitmessung von Korngrößenverteilungen durchgeführt. Dabei stellt sich eine Partial Least Squares Regression (partielle kleinste Fehlerquadrate Regression), angewandt auf geometrische Deskriptoren der Partikel, welche aus zweidimensionalen RGB-Bildern gewonnen wurden, als vielversprechender Ansatz heraus.

PUBLICATIONS AND (CO-)SUPERVISED THESES

In the following, publications and (co-)supervised students' theses elaborated in the context of the present doctoral thesis are presented.

Articles in peer-reviewed journals

Sarc, R., Curtis, A., Kandlbauer, L., **Khodier, K.**, Lorber, K.E., Pomberger, R., 2019. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management – A review. Waste management 95, 476–492. https://doi.org/10.1016/j.wasman.2019.06.035.

Khodier, K., Viczek, S.A., Curtis, A., Aldrian, A., O'Leary, P., Lehner, M., Sarc, R., 2020. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis. International journal of environmental science and technology 17 (2), 959–972. https://doi.org/10.1007/s13762-019-02526-w.

Möllnitz, S., **Khodier, K.**, Pomberger, R., Sarc, R., 2020. Grain size dependent distribution of different plastic types in coarse shredded mixed commercial and municipal waste. Waste management 103, *388*–398. https://doi.org/10.1016/j.wasman.2019.12.037.

Möllnitz, S., Küppers, B., Curtis, A., **Khodier, K.**, Sarc, R., 2021. Influence of pre-screening on down-stream processing for the production of plastic enriched fractions for recycling from mixed commercial and municipal waste. Waste management 119, 365–373. https://doi.org/10.1016/j.wasman.2020.10.007.

Kandlbauer, L., **Khodier, K.**, Ninevski, D., Sarc R., 2021. Sensor-based particle size determination of shredded mixed commercial waste based on two-dimensional images. Waste management 120, 784–794. https://doi.org/ 10.1016/j.wasman.2020.11.003.

Curtis A., Küppers B., Möllnitz, S., **Khodier K.**, Sarc R., 2021. Digital material flow monitoring in waste processing – the relevance of material and throughput fluctuations. Waste management 120, 687–697. http://doi.org/10.1016/j.wasman.2020.10.037.

Khodier, K., Feyerer, C., Möllnitz, S., Curtis, A., Sarc, R., 2021. Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: coarse-shredding of commercial waste. Waste management 121, 164–174. https://doi.org/10.1016/j.wasman.2020.12.015.

Viczek S.A., **Khodier, K.**, Kandlbauer, L., Aldrian, A., Redhammer, G., Tippelt, G., Sarc, R., 2021. The particle size-dependent distribution of chemical elements in mixed commercial waste and implications for enhancing SRF quality. Science of the total environment 776. https://doi.org/10.1016/j.scitotenv.2021.145343.

Khodier, K., Sarc, R., 2021. Distribution-independent empirical modeling of particle size distributions – coarse-shredding of mixed commercial waste. Processes 9 (3). https://doi.org/10.3390/pr9030414.

Book chapter

Sarc, R., Curtis, A., Kandlbauer, L., **Khodier, K.**, Lorber, K.E., Pomberger, R., 2020. Abfallwirtschaft 4.0 [Waste management 4.0], in: Frenz, W. (Ed.), Handbuch Industrie 4.0: Recht, Technik, Gesellschaft [Handbook industry 4.0: law, technology, society]. Springer, Berlin, Heidelberg, pp. 989–1014.

Conference contributions

Khodier, K., Sarc, R., Lehner, M., Pomberger, R., O'Leary, P., 2018. Process equipment for the implementation of industry 4.0-approaches on mechanical treatment of municipal solid waste, **poster and written contribution** in: Samhaber, W.M., Schneiderbauer, S., Rezaei, M. (Eds.), 14th Minisymposium chemical & process engineering and 5th particle forum: Book of Abstracts, p. 64. Minisymposium Verfahrenstechnik, Linz, Austria, 4/04/18.

Khodier, K., Sarc, R., 2018. Beschreibung von Abfallzusammensetzungen für Monte-Carlo-Simulationen – Ein Überblick über mathematische Möglichkeiten [Description of waste compositions for Monte Carlo simulations – an overview of potential mathematical methods], **written contribution** in: Pomberger, R. (Ed.), Recy & DepoTech 2018: VORTRÄGE-Konferenzband zur 14. Recy & DepoTech Band. vol. 14, pp. 799 - 804.

Sarc, R., Curtis, A., **Khodier, K.**, Koinegg, J., Ortner, M., 2018. Digitale Abfallwirtschaft [Digital waste management], **written contribution** in: Pomberger, R. (Ed.), Recy & DepoTech 2018: Vorträge-Konferenzband zur 14. Recy & DepoTech Band. vol. 14, pp. 793 - 798.

Khodier, K., Sarc, R., Lehner, M., O'Leary, P., 2019. Evaluating particle size distributions of solid waste – what is "good"?, **poster and written contribution** in: Conference Proceedings of 15th Minisymposium Verfahrenstechnik and 6th Partikelforum, Montanuniversitaet Leoben, Austria, 29th-30th April 2019. p. 85, 15th Minisymposium Verfahrenstechnik and 6th Partikelforum, Leoben, Austria, 29/04/19.

Möllnitz, S., **Khodier, K.**., Sarc, R., 2019. Kunststoffe im gemischten Gewerbeabfall – In welcher Form und Wo? [Plastics in mixed commercial waste – in what form and where?], **written contribution** in: Proceedings of Österreichische Abfallwirtschaftstagung 2019: Vom Wert- zum Werkstoff, Eisenstadt, Austria, 15/05/19–17/05/19.

Möllnitz, S., **Khodier, K.**, Sarc, R., 2019. Grain size dependent distribution of different plastic types in coarse-shredded mixed commercial waste, **written contribution** in: Proceedings of 10th International Symposium on Feedstock Recycling of Polymeric Materials (ISFR), Budapest, Hungary, 26/05/19–29/05/19.

Khodier, K., Lehner, M., Sarc, R. 2019. Multilinear modelling of particle size distributions, **poster and written contribution** in: Egozcue, J.J., Graffelman, J., Ortego, M.I. (Eds.), Proceedings of the 8th International Workshop on Compositional Data Analysis (CoDaWork2019), pp. 82–85, CoDaWork 2019, Terrassa, Spain, 03/06/19.

Möllnitz, S., **Khodier, K.**, Sarc, R., 2019. Grain size dependent distribution of plastic types in coarse-shredded mixed commercial waste, **written contribution** in: Proceedings of 7th International Conference on Sustainable Solid Waste Management, Heraklion, Crete Island, Greece, 26/06/19–29/06/19.

Viczek, S.A., **Khodier, K.**, Aldrian, A., Pomberger, R., Sarc, R., 2019. Grain size dependent distribution of contaminants in coarse-shredded commercial waste – results for As, Ba, Cd, Co, Cr, Cu, Hg, Mn, Mo, Ni, Pb, and Sb, **written contribution** in: Proceedings of 7th International Conference on Sustainable Solid Waste Management, Heraklion, Greece, 26/06/19–29/06/19.

Möllnitz, S., **Khodier, K.**, Sarc, R., 2019. Grain size dependent distribution of different plastic types in coarse-shredded mixed municipal solid waste, **written contribution** in: Proceedings of 17th International waste management and landfill symposium (sardinia_2019), Cagliari, Italy, 30/09/19–04/10/19.

Viczek, S.A., **Khodier, K.**, Pomberger, R., Sarc, R., 2019. Grain size dependent distribution of As, Cd, Cl, Co, Cr, Fe, Hg, Ni, Pb, Sb, Sn, Ti, V, W, and Zn in coarse-shredded commercial waste, **written contribution** in: Proceedings of 17th International waste management and landfill symposium (sardinia_2019), Cagliari, Italy, 30/09/19–04/10/19.

Khodier, K., Curtis, A., Sarc, R., Lehner, M., O'Leary, P., Pomberger, R., 2019. Smart solid waste processing plant: vision and pathway, **oral and written contribution** in: Proceedings of ISWA world congress 2019, Bilbao, Spain, 7/10/19–9/10/19.

Khodier, K., Sarc, R., Lehner, M., 2019. Result reliability in comparative solid waste processing experiments, **poster and written contribution** in: Proceedings of ISWA world congress 2019, Bilbao, Spain, 7/10/19–9/10/19.

Viczek, S.A., **Khodier, K.**, Pomberger, R., Sarc, R., 2019. Grain Size Dependent Distribution of Contaminants in Coarse-Shredded Commercial Waste – As, Cd, Co, Cr, Hg, Ni, Pb, Sb, and Cl, **written contribution (peer reviewed)** in: Proceedings of IRRC - Waste to Energy, 14/10/19–15/10/19 pp. 413–429.

Khodier, K., Lehner, M., Sarc R., 2020. Empirical modeling of compositions in chemical engineering, **oral and written contribution** in: 16th Minisymposium chemical & process engineering and 7th particle forum: Book of Abstracts. Minisymposium Verfahrenstechnik, Vienna, Austria, 22/09/20.

Khodier, K., Sarc, R., 2020. Digitalization and intelligent robotics in circular economy oriented waste management, **oral contribution** at: 11. Strokovna konferenca Okolje in odpadki, Ljubljana, Slovenia, 13/10/2020.

Khodier, K., Feyerer, C., Sarc, R. 2020. Optimierte Vorzerkleinerung gemischter Gewerbeabfälle auf Basis empirischer Modelle [Optimized coarse shredding of mixed commercial waste, based on empirical models], **oral and written contribution** in: Conference proceedings of the 15th Recy & DepoTech conference. Recy & DepoTech 2020, Leoben, Austria, 07/11/20–09/11/20.

Möllnitz, S**.**, Küppers, B., Curtis, A., **Khodier, K**., Sarc, R., 2020. Influence of pre-screening before ballistic separation on NIR-sorting quality of plastic rich 3D-fractions out of mixed wastes. **written contribution** in: Conference proceedings of the 15th Recy & DepoTech conference. Recy & DepoTech 2020, Leoben, Austria, 07/11/20–09/11/20.

Kandlbauer, L., Sarc, R., **Khodier, K.**, 2020. Charakterisierung von Partikeln gemischten Gewerbeabfalls über Partikeldeskriptoren zur sensorischen Messung der Korngröße [Characterisation of particles from mixed commercial waste by particle descriptors for sensor-based measurement of the particle size], **written contribution** in: Conference proceedings of the 15th Recy & DepoTech conference. Recy & DepoTech 2020, Leoben, Austria, 07/11/20–09/11/20.

(Co-)supervised Bachelor's and Master's Theses

Amann C., 2018. Magnettechnologien in der Abfallaufbereitung: Stand der Technik [Magnetic technologies in waste processing: state of the art], **Bachelor's Thesis**.

Köhldorfer P., 2018. Sensorgestützte Sortierung in der Abfallwirtschaft: Stand der Technik [Sensor-based sorting in waste management: state of the art], **Bachelor's Thesis**.

Wild, N.C., 2019. Aufbau eines Wartungs- und Störfalldokumentationssystems für die Maschinen einer mechanischen Abfallaufbereitung [Implementation of a maintenance and failure documentation for the machines in a mechanical waste treatment plant], **Master's Thesis**.

Kandlbauer, L., 2020. Sensorische Messung der Trommelsiebkorngrößenverteilung von gemischten Gewerbeabfällen [Sensor based measurement of the grain size distribution of mixed commercial waste according to drum screens], **Master's Thesis**.

Kodon, E., 2020. Modellierung der Trommelsiebklassierung: Effizienzsteigerung durch statistische Versuchsplanung [Modelling drum screen kinetics: Increasing the efficiency using the Design of Experiments], **Bachelor's Thesis**.

TABLE OF CONTENTS

Page

1   INTRODUCTION

In the European Union, around 5.2 metric tons of waste per capita were generated in 2018 (Eurostat, 2020) – as much as ten years before (Schrör, 2011), despite the priority of reduction and preparation for re-use, according to the waste hierarchy, which was first published in 2008 in the Waste Framework Directive (European Union, 2008). These significant amounts, which correspond to 2.3 billion metric tons in total in 2018, encourage the European Union's efforts concerning a transition towards a circular economy. Hence, the treatment and especially the recycling of waste has become a priority – which is demonstrated, for example, through the recycling targets set in the Circular Economy Package, requiring the recycling of 65% of municipal waste by 2035 (European Union, 2018).

Besides municipal waste, also considerable amounts of mixed commercial waste are produced each year, requiring efficient and effective treatment processes to enable its partial recycling and thermal exploitation. In Austria, it is classified under the waste code 91101 "municipal waste and similar commercial waste" (BMK, 2020), together with mixed municipal solid waste. According to Weißenbach et al. (2019), 1.1 million metric tons of mixed commercial waste were produced in Austria and 5.9 million metric tons in Germany in 2016, for example, leading to an increased awareness of their potential concerning the transition towards a circular economy. This awareness is demonstrated, for instance, by the German "Regulation on the management of commercial municipal waste and of certain construction and demolition waste" (German Federal Government, 2017), which demands the recycling of 30% of mixed commercial waste since 2019.

At the same time, the fourth industrial revolution, which is referenced with various terms like "industry 4.0", "internet of things," or "digitalization," is ongoing. According to Sarc et al. (2019), it describes the *"widespread introduction of information and communication technology (ICT) as well as its connection to an Internet of Things, Services and Data with the goal of real-time control of production and value chain networks"*. An obvious question that arises is what this fourth industrial revolution can contribute to the transition towards a circular economy. Even more, considering that the "digital readiness" of waste management companies in German-speaking countries is only about 30%, according to Berger (2016), as cited by Sarc et al. (2019).

Consequently, the Austrian Research Promotion Agency (Österreichische Forschungschungsförderungsgesellschaft, FFG) funded a four-year research program, with a total budget of around 4.9 million Euros, called "Recycling and Recovery of Waste 4.0" (ReWaste4.0), involving two research institutions and eight industry partners. It has started in April 2017 and is the framework within the research described in this thesis has been conducted. ReWaste4.0 targets the development and integration of digitalized approaches, con-

sidering the vision of smart future treatment plants, focusing on the mechanical treatment of non-hazardous mixed solid waste. It targets enhancing the efficiency of processing plants and the share of waste that is either recycled or thermally utilized.

1.1 Status quo

Mixed solid waste – commercial or municipal – is commonly first treated in mechanical processing plants (see, e.g., Müller and Bockreis, 2015). And coarse-shredding is usually the first step of this mechanical treatment (cf. Gundupalli et al., 2017; Möllnitz et al., 2021; Pomberger, 2008). Besides coarse shredders' role in terms of comminution and liberation of the particles, they are the primary dosing devices of most plants, according to Feil and Pretz (2018), and hence a major influence factor on the overall subsequent process behavior. Considering that screens are also commonly used machines in mechanical waste processing plants, the shredder's influence on the particle size distribution of the waste moreover affects the path particles take through the plant, and thereby also the machines and related process steps they pass through.

Despite its importance for waste treatment, there is only little literature on coarse-shredding of mixed solid waste. Existing studies, include the works of Luo et al. (2010) and Luo et al. (2011): they examine the performance of a novel shredder in a lab-scale, using prepared model material, focusing on the comminution of organic fractions – hence, their results cannot be transferred to industry-scale shredding of mixed solid waste.

Zhang et al. (2019) study the particle size distributions produced by a variety of comminution machines for mixed municipal solid waste at different machine settings. But since they analyze the waste at the interface from mechanical to biological treatment, and consequently after more processing steps than just shredding, no conclusions on the distribution of the shredded, unsorted waste can be drawn. Furthermore, while they repeated some measurements, no statistical analyses of the results were performed.

Feil and Pretz (2018) compare the throughput steadiness of a single-shaft and a two-shaft coarse shredder in a standard operation, discussing its impact on the performance of subsequent machines. But they do not vary machine parameters, and hence do not discuss the optimality of their settings.

Eventually, to the best of the doctoral candidate's knowledge, no quantitative studies that investigate the influence of shredder parameters on their real-scale behavior, proving the results either by explaining them with physical models, or based on statistical tests, have been published yet.

Interviews with experts from industry showed, that also in industrial performance tests for machine purchases, machines are mostly compared based on limited data, lacking statistical

validity. And the machines are tested with fixed parameter settings: for coarse shredders, the radial gap is usually completely closed, while the shaft rotation speed is set to the maximum – a setting whose optimality needs to be questioned, based on the results of this thesis.

Besides machine tests, coarse shredders are also operated statically in operating plants. In practice, shaft rotation speeds are adjusted by hand, from time to time, and the programming of the shredders incorporates some dynamic behavior – like reversing in regular time intervals for cleaning or when some particle gets stuck. Nonetheless, the shredders do not react to the waste, despite its influence on the throughput and the plant's overall performance. Hence, a non-controlled, highly significant unsteadiness of the processes, as visible in the measurements of Curtis et al. (2021), for example, is accepted in the absence of counteracting strategies.

Physical numerical models of mechanical processing of particulate materials are usually based on the Discrete Element Method (DEM). Published examples include the works of Pieper et al. (2016) on optical belt sorting and Lee et al. (2008) on the impact breakage of concrete waste. But at present, the models in literature are not capable of reflecting the variability of mixed commercial waste and the resulting complex material-material interactions and material-machine interactions. And it remains questionable whether physical models, which require detailed data on the geometry and the properties of the processed particles, are a promising choice for investigating the industry-scale behavior of processing machines for mixed solid waste. Hence, the focus and novelty of the present thesis is the empirical modeling of real-scale coarse shredding, including the produced non-scalar particle size distributions, using polynomial regression models.

1.2 Problem formulation

While regression modeling is a standard procedure, at first glance, doing so turns out to be quite complex in detail in the context of mixed solid waste shredding. Generating the data involves industry-scale experiments, in which the highly variable feed – mixed commercial waste – adds significant distortion to the data (Khodier et al., 2020). And the extent to which this distortion can be counteracted by a high number of experimental runs is limited due to the high costs of such experiments.

Furthermore, concerning analyses of the product material, sampling is involved, since in each experiment in this thesis up to about 40 metric tons of waste were processed. While there are sampling standards, which provide advice, these often deviate from sampling theory for practical purposes, without reporting expectable magnitudes of introduced sampling errors (e.g., ÖNORM S 2127: Austrian Standards Institute, 2011). Therefore, they are not immediately applicable to scientific purposes, but rather an in-depth understanding of sampling theory is needed to enable a critical evaluation of the standards' directions.

Finally, modeling scalar process properties, like throughput or energy demand, can be done, applying standard (despite not necessarily simple) statistics. Modeling particle size distributions, on the other hand, is far from trivial: depending on the chosen descriptive technique, these are either continuous functions or compositional vectors of particle size classes – constrained to a vector space, called the simplex (cf. Pawlowsky-Glahn et al., 2015), which is represented by the well-known ternary diagram for three dimensions. As a result, the validity of many standard statistical procedures is harmed. Consequently, how to model the produced particle size distributions empirically is one of the research questions of this thesis.

2 CONCEPT OF THE THESIS

The targeted real-time process control with the potential inclusion of smart algorithms requires three essential elements (see Figure 1), as discussed by Khodier et al. (2019): real-time metrology, for measuring the actual state of the system to be controlled (e.g., throughput, particle size distribution), actuators that allow a controller to change the process somehow (physical actuators like rotation speed of a drum screen, as well as software actuators like a classification threshold of a sensor-based sorter), and control algorithms, that control the change of actuators to improve a target value, that is based on the data from metrology, but also potentially other data, like actual market prices.
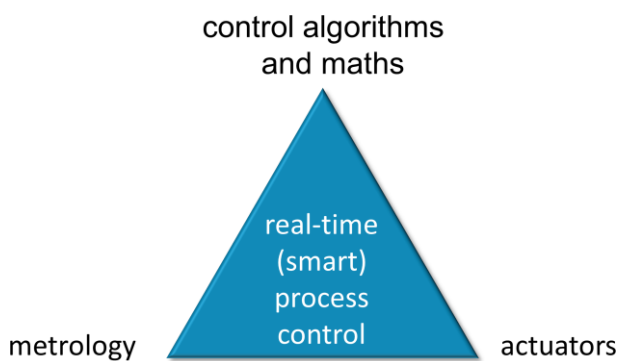
Figure 1: Trinity of real-time process control

Classical or smart control algorithms are already being applied in different industries and are therefore state of the art. But many potential actuators in solid waste processing plants are not digitally controllable (if at all) nowadays due to the lack of knowledge on their influence and a lack of economical usefulness of doing so, considering the state of the art. Furthermore, the plants are only poorly equipped with metrology, in particular regarding material quality, due to technical and economical limitations.

2.1 Objective

This thesis aims to gather, develop, and document empirical modeling methodologies that allow drawing significant and therefore trustworthy conclusions concerning machine parameter influences on solid waste processing, despite the disturbance caused by the heterogeneity of waste. It furthermore targets contributing to the online measurability of particle-size distributions in real-time. Concerning parameter influences, it particularly examines their kind and magnitude regarding coarse-shredding of mixed commercial waste, with shredding being the most sophisticated mechanical treatment step to investigate, since the shredded material cannot be re-used in the experiments. Concretely, the influences of the radial shredder gap width and the shaft rotation speed and – to provide a meaningful relation – also the cutting

tool geometry, on the throughput behavior (total and steadiness), energy demand, and on the produced particle size distributions, is examined.

2.2 Research gap

While experience and engineering intuition provide rough estimates concerning the influences of the parameters above (see section 2.1), no quantitative information on their impact on process parameters and material qualities is available. Furthermore, published literature on waste comminution lacks documenting techniques to draw statistically significant conclusions on parameter influences from real-scale experiments with mixed solid waste. Additionally, modeling influences on particle size distributions poses mathematical challenges. These distributions are describable as continuous probability density functions, or more usual, as compositions of discrete particle size classes, covering different particle size ranges. Hence, the modeled dependent variables are non-scalar with interdependent dimensions. Thus, methods for considering their special properties need to be found.

Furthermore, sensor-based online measurement of particle size distributions is only performed based on counting pixels of the projected area, without questioning the relation of the determined particle sizes to particles' behavior in screens. Hence, methods that reflect the particle sizes, according to screening, need to be found.

2.3 Research questions

Considering the objective of this thesis, the discussed challenges, and the status quo concerning modeling mechanical processing of mixed commercial waste, the following research questions were defined:

1. How can the processing products be representatively sampled in coarse-shredding experiments with mixed commercial waste, and how large are the remaining sampling errors?
2. How can reliable conclusions about machine influences on waste shredding be drawn, despite sampling errors and inter-experimental differences in the waste?
3. How can particle size distributions be empirically modeled?
4. How big is the influence of the radial gap width, the shaft rotation speed, and the cutting tool geometry of an industry-scale coarse shredder on the throughput behavior, the energy demand, and the particle size distribution of shredded mixed commercial waste?
5. How can the particle size distribution of coarsely shredded mixed commercial waste, according to a screen, be determined in real-time, using state-of-the-art two-dimensional sensor data?

2.4 Methodology

Methodologically, the work in this thesis is based on three major elements: research on the state of the art, research on and application of mathematical methods for experimentation and analysis, and planning and execution of real-scale experiments for data acquisition.

2.4.1   Research on the state of the art

Before conducting own practical research, it is indispensable to gain a comprehensive overview of the state of the art and science. Hence, extensive research was carried out to gather knowledge on the state of mechanical waste processing in general, digitalization in the waste industry in particular, and modeling of mechanical processing of mixed solid waste. This knowledge was achieved through literature and patent research, expert interviews, planning and leading a workshop in cooperation with the Styrian Green Tech Cluster, and visiting fairs and industrial plants.

2.4.2   Research on mathematical methods

The efficient and reliable performance and analysis of waste shredding experiments need a deep understanding of many basic as well as more exotic mathematical and statistical concepts. Within this thesis, comprehensive knowledge on especially the following fields was established to enable proper and profound planning, execution, and analysis of the performed experiments: theory of sampling, design and analysis of experiments, univariate and multivariate multiple linear regression, including (multivariate) analysis of variance, as well as modeling and analysis of compositional data.

2.4.3   Industry-scale experiments

In many research fields – including chemical and process engineering – it is usual practice to first experiment in lab-scale, then move on to pilot-scale, and finally progress to industry-scale. While this is resource-efficient and facilitates fundamental research, it is hardly applicable for mixed solid waste processing, mainly for one reason: the feed material (the waste) cannot properly be down-scaled. Hence, to understand mechanical processing (and especially shredding) of real waste, the experiments must be executed in real-scale. Therefore, numerous industry-scale experiments were conducted within this thesis, mostly using industry-scale mobile machinery and real waste. The core experiment of this thesis is a 32 runs experiment on the shredding of mixed commercial waste, including material analysis. Due to its scale, it consumed enormous resources in terms of machinery (occupation of six industry-scale machines for many weeks) and manpower (approximately 3500 working hours in total) while shredding about 700 metric tons of waste. Consequently, it required thorough project

and resources management, including machine and staff availability, staff training, as well as the clarification of legal requirements for that scale of waste processing.

2.5 Structure of the thesis

The overall structure of the thesis is shown in Table 1. It is divided into three main parts. In the first part, "Introduction and concept," a general introduction into the field of research and the topic of the thesis is given. The ensuing chapter presents the concept of this work, including the presentation of the research gap and the research questions, and the structure of the document.

Table 1: Structure of the thesis (Ch. 1–Ch. 5 stand for the chapters of the thesis and P I–P IV stand for the research papers)

Introduction and concept	Ch. 1	Introduction	
	Ch. 2	Concept of the thesis	
Practical part	Ch. 3	Results	Empirical modeling and shredder parametrization (**P I**, **P II**, **P III**)
			Online particle size measurement (**P IV**)
Conclusive part	Ch. 4	Summary and discussion	
	Ch. 5	Conclusions and outlook	

The second part of the thesis is the practical part, which consists of the documentation of performed research activities in research papers. It is subdivided into two sections. The first, "Empirical modeling and shredder parametrization," contains the core research activities of the doctoral candidate. These include an extensive study on sampling quality (**publication I**): the theory of sampling is explained and applied on coarsely shredded mixed commercial waste, followed by a practical evaluation of the general estimation error caused by sampling and analytics, determined through a so-called replication experiment. Ensuing, a research paper on the derivation of significant results from mechanical waste processing experiments is presented (**publication II**), covering the question of how to extract trustworthy conclusions on the behavior of scalar dependent variables, despite the variability of the waste used in the experiments. Concretely, it investigates the influence of the radial gap width, the shaft rotation speed, and the cutting tool geometry of coarse shredders on their throughput, throughput steadiness, and specific energy demand. The last paper of this section (**publication III**) deals with extending the methods presented in publication II on multivariate compositional dependent variables, more precisely: particle size distributions, which are described by a set of discrete particle size classes, determined through sieve analyses. It applies the methods to the question of how the shredders' parametrization influences the quality of the shredding product in terms of the total particle size distribution.

The second section of the practical part, "Online particle size measurement," deals with the real-time measurement of the particle size distributions of coarsely shredded mixed commercial waste, according to a drum screen, from visual and near-infrared sensor-data of single particles. This approach aims at providing the first online method for particle size measurement of such waste (**publication IV**). The allocation of the four described research papers to the different elements of the trinity of real-time process control (as shown in Figure 1) is shown in Figure 3.
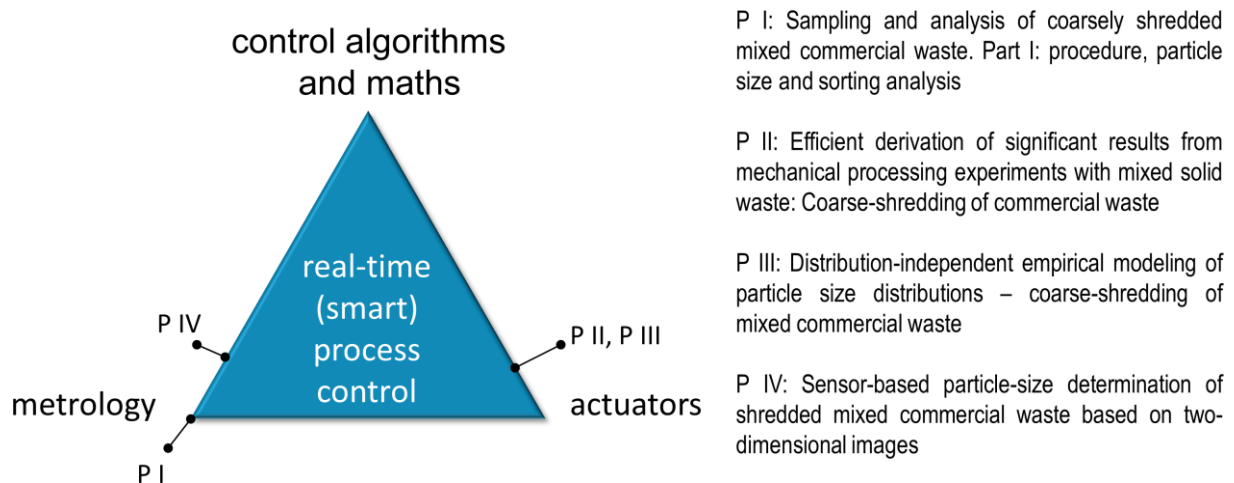
Figure 2: Allocation of the publications to the elements of the trinity of real-time process control

In the "Conclusive part" of the thesis, the research questions are answered and discussed, based on the publications at hand, in the chapter "Summary and discussion". In the final chapter, conclusions are drawn, and suggestions for further research and an outlook on future research activities are given.

3 RESULTS

3.1 Empirical modeling and shredder parametrization

Publication I

Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis

K. Khodier, S.A. Viczek, A. Curtis, A. Aldrian, P. O'Leary, M. Lehner, R. Sarc

International Journal of Environmental Science and Technology, 2020, 17 (2), 959–972, https://doi.org/10.1007/s13762-019-02526-w

Author contributions according to the CRediT system:

KK: conceptualization, methodology, software, formal analysis, investigation, data curation, writing – original draft preparation, writing – review & editing, visualization, project administration

SV: conceptualization, methodology, investigation, writing – review & editing

AC: methodology, investigation

AA: methodology

PO: supervision

ML: writing – review & editing, supervision

RS: methodology, investigation, writing – review & editing, supervision, funding acquisition

ORIGINAL PAPER

Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis

K. Khodier[1] · S. A. Viczek[2] · A. Curtis[2] · A. Aldrian[2] · P. O'Leary[3] · M. Lehner[1] · R. Sarc[2]

Abstract
Performing experiments with mixed commercial waste, sampling is unavoidable for material analysis. Thus, the procedure of sampling needs to be defined in a way that guarantees sufficient accuracy regarding the estimation of the examined analytes. In this work, a sampling procedure for coarsely shredded mixed commercial waste, based on the Austrian Standard ÖNORM S 2127, the horizontal sampling standard DS 3077 and the theory of sampling, was established, described and examined through a replication experiment determining the relative sampling variability. The analytes are described through a matrix of nine (9) material classes and nine (9) particle size classes. It turns out that the typical threshold value of 20% can be reached for some fractions of the particle size–material matrix (for example, wood 20–40 mm and cardboard 60–80 mm) but gets as bad as 231% (wood 200–400 mm) for others. Furthermore, a decrease in the relative sampling variability with the mass share of a fraction is observed. Part of the observed variability is explainable through the fundamental sampling error, while contributions of other types of sampling errors are also evident. The results can be used for estimating confidence intervals for experimental outcomes as well as assessing required sample sizes for reaching a target precision when working with mixed commercial waste.

List of symbols

$\mathbf{A}_v$	Binary matrix for combining adjacents of v particle size fractions [–]
$\mathbf{B}_w$	Binary matrix for combining 1 to w material classes [–]
c	Constitutional parameter [kg/m³]
CV	Coefficient of variance [–]

✉ R. Sarc
renato.sarc@unileoben.ac.at

1 Chair of Process Technology and Industrial Environmental Protection, Department of Environmental and Energy Process Engineering, Montanuniversitaet Leoben, Franz-Josef-Strasse 18, 8700 Leoben, Austria

2 Chair of Waste Processing Technology and Waste Management, Department of Environmental and Energy Process Engineering, Montanuniversitaet Leoben, Franz-Josef-Strasse 18, 8700 Leoben, Austria

3 Chair of Automation, Department Product Engineering, Montanuniversitaet Leoben, Peter-Tunner-Strasse 25/II, 8700 Leoben, Austria

d_{05}	5th percentile particle size [mm], [cm]
d_{95}	95th percentile particle size [cm]
d_{max}	Maximum particle diameter [mm]
f	Particle shape parameter [–]
f_{red}	Mass reduction factor [–]
f_{red}^*	Real mass reduction factor [–]
$f_{red,r}^*$	f_{red}^* when reducing the fine fraction of the $(r-1)$th screening step [–]
g	Particle size parameter [–]
HI_{lot}	Heterogeneity invariant of the lot [g]
$\mathbf{M}$	Particle size–material matrix (masses) [kg]
m_c	Average particle mass of the constituent c [g]
m_{disc}	Mass discarded during mass reduction [kg]
m_{ij}	Mass of the ith particle size fraction and jth material class in the primary sample [kg]
m_{ij}^*	Weighed mass of the ith particle size fraction and jth material class [kg]
m_{inc}	Minimum increment mass [kg]
m_k	Average particle masses of the constituents k [g]
m_{lot}	Mass of the lot to be sampled [g]
m_{part}	Mass of sample part [kg]
m_{pres}	Mass preserved during mass reduction [kg]

Springer

m_{sam}	Minimum sample mass [g]
$m_{\text{sam},k}$	Mass of the kth sample [kg]
n	Number of samples [–]
p	Fraction of particles with a specific characteristic [–]
q	Number of constituents [–]
RSV	Relative sampling variability [%]
s	Standard deviation [conc. u.]
t_{inc}	Increment extraction time [s]
$\dot{V}$	Volume flow [kg/m^3]
v	Number of particle size fractions [–]
$\mathbf{W}$	Particle size-material matrix (mass shares) [kg/kg]
$\mathbf{W}_{\text{mp}}$	Matrix of all regarded particle size fraction and material class combinations (mass shares) [kg/kg]
w	Number of material classes [–]
w_c	Mass share of the constituent c [g]
w_{ij}	Mass share of the ith particle size fraction and jth material class in the primary sample [kg/kg]
w_k	Mass shares of the constituents k [g]
$\bar{x}$	Weighted arithmetic mean [conc. u.]
x_k	Concentration of a specific analyte according to sample k [conc. u.]
β	Liberation parameter [–]
ρ	Bulk density [kg/m^3]
σ_{FSE}^2	Variance caused by the fundamental sampling error [–]

Introduction

Coarse shredding followed by one or more screening stages is often the first step when processing mixed commercial waste (MCW). Besides size reduction and definition, this combination contributes to the concentration of different materials as well as their contained chemical elements in different fractions. Reasons for this are the different particle size distributions of the material classes in the original material, as well as differences in comminution behaviour—e.g. brittle fracturing of glass and passing through or tearing of plastic foils.

Aiming at systematically steering this concentration process, optimal shredder and screen parametrization are intended to be found through empirical regression models for the particle size–material matrix as well as for the distribution of the concentrations of contained elements over particle sizes, based on experimental results. Because of the high inherent inhomogeneity of MCW, performing such experiments demands processing of large amounts of material to homogenize the variability of the input stream regarding composition and particle size between the single

runs. These high amounts of material cause the infeasibility of analysing the complete shredding product so that samples need to be taken.

Various standards and recommendations concerning sampling of MCW are available, giving guidelines on sample extraction, increment sizes, sample numbers and sizes, and sample processing for different applications. Regarding masses, they aim at ensuring that the amount of analyte contained in the sample is sufficient to keep the effect of single particles ending up in the sample—or not—at an insignificant level. This is, for example, done through average particle masses and shares of sorting fractions by Felsenstein and Spangl (2017). In contrast, the technical report CEN/TR 15310 (European Committee for Standardization 2006) for uses inhomogeneity descriptors in combination with the cubic diameter of the largest particles, as well as bulk density, as an estimate for maximum particle contributions to the analyte.

In Austria, the Austrian Standard ÖNORM S 2127— Basic characterization of waste heaps or solid waste from containers and transport vehicles (Austrian Standards Institute 2011)—is usually applied, demanding a minimum increment mass m_{inc} [kg] according to Eq. (1)—where d_{95} is the 95th percentile particle size [mm]—and a minimum number of 10 increments per (representative) sample. Furthermore, at least one sample per 200 t of waste investigated is required. Multiplying these requirements leads to a minimum for the total sample mass.

$$m_{\text{inc [kg]}} \geq 0.06 \cdot d_{95[\text{mm}]} \tag{1}$$

However, the standard does not give information about the statistical significance of the sampling result. Furthermore, the linear consideration of the diameter for the resulting sample mass is very likely to underestimate the masses required to get reliable information about coarse fractions. It is a compromise between reliability and practicability in terms of comprehensibility and economically feasible (sorting) analyses of the resulting sample masses. According to Wavrer (2018), a comparable practically oriented approach called MODECOM™ is used in France.

The technical report CEN/TR 15310—characterization of waste—sampling of waste materials (European Committee for Standardization 2006) is another available reference. It defines the minimum increment mass m_{inc} [kg] according to Eq. (2), where ρ [kg/m^3] is the bulk density of the material.

$$m_{\text{inc [kg]}} \geq 2.7 \cdot 10^{-8} \cdot \rho_{\left[\frac{\text{kg}}{\text{m}^3}\right]} \cdot d_{95\,[\text{mm}]} \tag{2}$$

The equation describes the resulting mass, when using a sampling device which is at least three times as long as the

maximum particle diameter (practically determined through d_{95}) in each dimension, as demanded by the report. This shall ensure that all particles can easily enter the sampling device. The resulting increment mass is about 50 times the maximum particle mass, according to the report. It furthermore suggests a minimum sample mass m_{sam} [g], calculated according to Eq. (3), where p [m/m] is the fraction of the particles with a specific characteristic, g [–] is the correction factor for the particle size distribution of the material to be sampled, and CV [–] is the desired coefficient of variation caused by the fundamental error.

$$m_{\mathrm{sam}[\mathrm{g}]} \geq \frac{1}{6} \cdot \pi \cdot \left(d_{95\,[\mathrm{cm}]}\right)^3 \cdot \rho_{\left[\frac{\mathrm{g}}{\mathrm{cm}}\right]} \cdot g_{[-]} \cdot \frac{\left(1 - p_{[\mathrm{m/m}]}\right)}{\mathrm{CV}^2_{[-]} \cdot p_{[\mathrm{m/m}]}} \tag{3}$$

The value of g depends on the quotient of the 95th and 5th percentile particle sizes (d_{95}/d_{05}), according to Eq. (4). 0.1 is suggested as a well-accepted value for CV, and p needs to be determined from knowledge about waste consistency.

$$g = \begin{cases} 0.25 & \text{if } 4 < d_{95}/d_{05} \\ 0.50 & \text{if } 2 < d_{95}/d_{05} \leq 4 \\ 0.75 & \text{if } 1 < d_{95}/d_{05} \leq 2 \\ 1 & \text{if } d_{95}/d_{05} = 1 \end{cases} \tag{4}$$

While the report also provides formulae for calculating the significance of the analytical results, determining it requires a priori knowledge about the material, as does the determination of minimum increment and sample masses.

Another reference is the Guidelines for statistical evaluation of sorting and particle gravimetric analyses from Vienna University of Technology (TU Wien) and the University of Natural Resources and Life Sciences in Vienna (BOKU) (Felsenstein and Spangl 2017). It provides theory-based instructions on required sample masses for different significance levels, while mainly addressing the sampling of the total mixed municipal waste of Austrian federal provinces and demanding prior particle weight analyses for calculating sample masses.

Gy (2004a), the founder of the theory of sampling (TOS), also provides formulae supporting the determination of required sample masses by calculating the minimum possible sampling error—the fundamental sampling error (FSE). It is calculated according to Eq. (5), where σ^2_{FSE} [–] is the variance caused by the FSE, m_{sam} [g] and m_{lot} [g] are the masses of the sample and of the lot to be sampled and $\mathrm{HI}_{\mathrm{lot}}$ [g] is the heterogeneity invariant of the lot (Gy 2004a). $\mathrm{HI}_{\mathrm{lot}}$ can be calculated through Eq. (6), which contains the following parameters:

- c [g/cm³]: constitutional parameter, which can vary from values lower than 1, up to millions
- β [–]: liberation parameter with $0 \leq \beta \leq 1$
- f [–]: particle shape parameter with $0 \leq f \leq 1$ and common values near 0.5
- g [–]: size range parameter with $0 \leq g \leq 1$
- d_{95} [cm]: 95th percentile particle size

$$\sigma^2_{\mathrm{FSE}[-]} = \left(\frac{1}{m_{\mathrm{sam}[\mathrm{g}]}} - \frac{1}{m_{\mathrm{lot}[\mathrm{g}]}}\right) \cdot \mathrm{HI}_{\mathrm{lot}[\mathrm{g}]} \tag{5}$$

$$\mathrm{HI}_{\mathrm{lot}[\mathrm{g}]} = c_{\left[\frac{\mathrm{g}}{\mathrm{cm}^3}\right]} \cdot \beta_{[-]} \cdot f_{[-]} \cdot g_{[-]} \cdot d^3_{95\,[\mathrm{cm}]} \tag{6}$$

While Gy states that values for these parameters can be found for different materials in literature, no references were found for MCW. Furthermore, according to Gy (2004a), no satisfactory formula is known yet for determining f. Hence, the heterogeneity invariant needs to be evaluated experimentally.

However, according to Wavrer (2018), a simplified formula exists for "simple particles," meaning cases where the particles are assumed to consist either of 0% or 100% of an analyte—as is typically the case for waste sorting analyses, where each particle is assigned to a sorting fraction. In that case, the FSE can be calculated through Eq. (7), where m_c and m_k are the average particle masses of the constituent of interest c and the other constituents k, and w_c and w_k are their mass shares, respectively. q stands for the number of constituents.

$$\begin{aligned} \sigma^2_{\mathrm{FSE}[-]} = & \left(\frac{1}{m_{\mathrm{sam}[\mathrm{g}]}} - \frac{1}{m_{\mathrm{lot}[\mathrm{g}]}}\right) \\ & \cdot \left(m_{c[\mathrm{g}]} \frac{1 - 2\,w_{c[-]}}{w_{c[-]}} + \sum_{k=1}^{q} w_{k[-]} \cdot m_{k[\mathrm{g}]}\right) \end{aligned} \tag{7}$$

Still, a priori knowledge is needed in terms of average particle masses and assumptions about the composition of the constituents, as is the case in Felsenstein and Spangl's (2017) guideline. Moreover, applying the formulae for the FSE does not provide information about the real sampling error beyond the contribution of the fundamental one.

In conclusion, no satisfactory guidance was found for a priori determination of required increment masses or sample masses for achieving a certain level of significance when sampling MCW. Furthermore, the general estimation error (GEE) for the elements of the particle size–material

Fig. 1 Contributions to the general estimation error
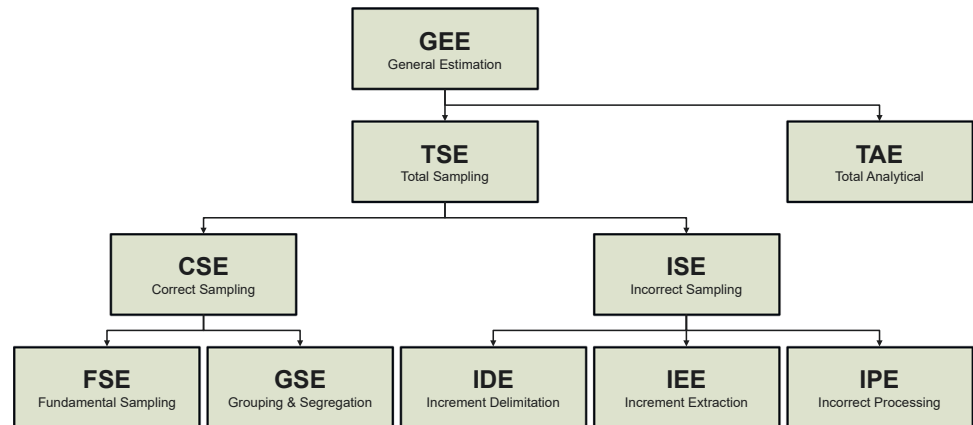
matrix as well as for the distribution of chemical elements throughout particle sizes is the result of several processing steps and subsampling steps. Thus, to evaluate analytical data quality for modelling purposes, as well as for interpreting experiments, the total error of the data acquisition process, from primary sampling to chemical analysis (which is the GEE), needs to be determined experimentally. This is done through a replication experiment (REx) as described in the Danish standard DS 3077, which is a horizontal sampling standard based on the TOS (Danish Standards Foundation 2013).

Even though the other described references may deliver more profound statements regarding necessary sample masses, the Austrian standard ÖNORM S 2127 was chosen for the REx in this work for multiple reasons: the necessary information about the material (which is only d_{95}) was available and analysing the resulting masses was expected to be feasible in practice. Furthermore, the REx offered the opportunity to evaluate this standard, which is widely applied in Austria.

The investigation to be presented will be published in two parts. In part I (i.e. the present contribution), a procedure for sampling MCW-shredding experiments is developed, based on TOS, DS 3077 and ÖNORM S 2127. Furthermore, the applied steps of sample processing from the primary sample to the particle size–material matrix are described. Finally, the results of a REx are presented, providing information about data quality when applying the described procedure. The corresponding experimental work was conducted from October to December 2018 in Allerheiligen im Mürztal, Styria, Austria.

Part II, presented by Viczek et al. (2019), deals with the distribution of several chemical elements, especially heavy metals, in different grain size fractions of coarsely shredded MCW. Post-sorting processing of the material for analysing the concentrations of these elements is described. Furthermore, the GEE is evaluated for the concentrations in different particle size classes through the REx. Ultimately, the distribution of the chemical elements throughout particle sizes, as well as correlations to the results of sorting analysis are presented.

Materials and methods

Theory of sampling

According to Esbensen and Wagner (2014), the theory of sampling is a universal, scale-invariant fundamentum for understanding sampling and the potential errors caused by it. It is based on the fundamental sampling principle, requiring all increments of the lot to have the same likelihood of ending up in the (representative) sample. It describes all errors contributing to the total sampling error (TSE). The GEE consists of this TSE in addition to the, often well-determined, total analytical error (TAE) (Esbensen and Wagner 2014). Its components are shown in Fig. 1.

Gy (2004b) divides the contributions to the TSE into correct sampling errors (CSE) and incorrect sampling errors (ISE). The first are caused by constitutional and distributional heterogeneities of the material to be sampled and are unavoidable, whereas the latter come from avoidable sampling mistakes and should, therefore, be avoided as far as possible.

CSE consists of two kinds of errors, the FSE and the grouping and segregation error (GSE). The FSE is caused by the constitutional heterogeneity, which describes chemical and physical differences between fragments of the lot—in the case of MCW: particles—and can only be altered

through physical interventions like comminution. The GSE, on the other hand, is present due to the distributional heterogeneity, meaning the spatial distribution of different particles, e.g. through segregation, or regarding MCW because of compositional differences of different, joined but not homogenized, waste sources (Esbensen and Julius 2009). The distributional heterogeneity is the reason why samples need to consist of a number of increments spaced around the lot to achieve acceptable levels of GSE.

ISEs are the sum of incorrect delimitation errors (IDEs), incorrect extraction errors (IEEs) and incorrect processing errors (IPEs). The IDEs describe errors in defining geometrical domains to be potentially taken as a sample. According to Gy (2004b), they can be avoided when collecting materials of a stream using equal time intervals. The IEEs appear, when the delimited domain (including all particles whose centre of mass is contained) cannot be precisely extracted, meaning particles end up in the sample that should not have and vice versa. Finally, IPEs mean all errors caused by incorrect processing of the sample after extraction and consist of six elements: contamination by foreign material, loss of material (e.g. dust), alteration in chemical and alteration in physical composition, involuntary operator faults and deliberate faults for manipulating results (Gy 2004b).

Replication experiment

The GEE manifested as the variability of repeated sampling can be quantified through a REx as described in DS 3077 (Danish Standards Foundation 2013). It is performed by extracting and analysing replicate samples of the same lot. The variability in analytical results obtained for these repeated samples is then expressed through the relative sampling variability (RSV), giving a measure for sampling quality evaluation. It is calculated according to Eq. (8), where s [arbitrary concentration unit (conc. u.)] is the standard deviation and $\bar{x}$ [conc. u.] is the arithmetic mean of the concentrations of a specific analyte in the repeated samples—functioning as an estimate for the true value.

$$\text{RSV}_{[\%]} = \frac{s}{\bar{x}} \cdot 100_{[\%]} \tag{8}$$

In this investigation, the single samples cover equal time spans in which material falls from the conveyor belt. Hence, for calculating $\bar{x}$, the contribution of each sample needs to be weighted by sample mass, as time spans with lower throughputs contribute less to the concentrations in the total lot. Therefore, $\bar{x}$ is calculated according to Eq. (9), where $m_{\text{sam},k}$ is the sample mass [kg] of the kth

replicate sample and x_k [conc. u.] is the concentration of a specific analyte according to sample k. As each of the samples is equally likely to be taken, s is calculated without weighting, according to Eq. (10), where n [–] is the number of samples.

$$\bar{x} = \frac{\sum_k x_k \cdot m_{\text{sam},k}}{\sum_k m_{\text{sam},k}} \tag{9}$$

$$s = \sqrt{\frac{1}{n-1} \cdot \sum_k \left(x_k - \bar{x}\right)^2} \tag{10}$$

According to DS 3077, the absolute minimum number of replicates for performing a REx is 10 (Danish Standards Foundation 2013). Due to the enormous amount of manual work needed in waste sorting analytics, this minimum number of samples is chosen for this work. Furthermore, the standard gives guidance for RSV interpretation, stating that 20% is a consensus acceptance threshold. This value is to be understood as a rough indication—the threshold applied in practice needs to be defined based on the potential impacts of analytical uncertainties.

Shredding experiment and primary sampling

Experimental set-up

The shredding experiment in this investigation was performed using a mobile single-shaft coarse shredder Terminator 5000 SD with the F-type cutting unit from the Austrian company Komptech (Fig. 2). It was fed using an ordinary wheel loader. The shredding product was discharged using the conveyor belt included in the machine, forming a windrow.

The experiment was performed operating the machine on 60% of the maximum shaft rotation speed (18.6 rpm) and with the cutting gap completely closed. The waste used was MCW from Styria in Austria.

Sample and increment mass

For defining the total sample mass to be taken, in this work the Austrian standard ÖNORM S 2127 was considered as a reference. Multiplying the minimum increment mass from Eq. (1) with the minimum number of increments, which is 10, and considering that the expected amount of waste to be processed in the experiment is less than 200 t, the minimum sample mass m_{sam} is calculated according to Eq. (11). With

400 mm being a conservative estimate for d_{95} from prior experiments, a minimum sample mass of 240 kg is defined.

$$m_{sam\,[kg]} \geq 0.6 \cdot d_{95[mm]} \tag{11}$$

As a rising number of increments forming a sample of a mass m_{sam} leads to better spatial coverage of the lot, it is very likely that dividing this mass into more than 10 increments—resulting in values for m_{inc} lower than defined by Eq. (1)—does not negatively influence sampling quality, but rather improve it while keeping the total mass to be analysed constant. For practical reasons consisting of the manageable sample device volume and increment mass when sampling by hand, as well as the maximum practicable sampling frequency and the target of keeping the experimental duration as short as possible because of the high throughput, 20 was chosen as the number of increments, resulting in a minimum increment mass of 12 kg.

Practical implementation of primary sampling

For reasons of practical implementation, sampling during the shredding experiments was performed by hand. This was done by holding a suitable open container into the falling stream at the end of the product conveyor belt, allowing preferable one-dimensional sampling. The container was held by two people standing at each side of the conveyor belt. At certain times, they received a starting signal to introduce the container into the stream. After a defined time (determination described below), the container was removed, containing one increment. To guarantee the accessibility of the belt, it was kept low during sampling intervals, needing the shredder to keep moving in the opposite direction of the output material stream, forming a long windrow.

According to CEN/TR 15310, each dimension of the sampling device should be at least three times as long as d_{95}, to allow the entry of all particles (European Committee for Standardization 2006). For the described sampling method, a container of $(1.2\ m)^3$ would have been unmanageable. The inner dimensions of the sampling device used (built from two mortar buckets) are $1.17 \times 0.37 \times 0.30$

Fig. 2 Feeding of Komptech Terminator 5000 SD

(length × width × depth in m), which corresponds to a volume of 0.13 m³. With the width of the conveyor belt being 1 m and holding the device very close to the belt, it was observed that all falling material entered the container.

For determining the duration of each sampling step, the mass flow was estimated at the beginning of the experiment. To do so, the mean volume flow on the conveyor belt was measured for a duration of 3 min using a laser triangulation measurement bar above the end of the belt. The REx was part of an experimental series: prior to the experiment, a calibration experiment was performed, processing a total mass of 3.5 t for linking volume flows to mass flows. It showed a bulk density of 161.8 kg/m³. At the end of the three minutes, the sampling duration for the extraction of an increment, t_{inc} was calculated from the average volume flow $\dot{V}$, the bulk density estimate ρ and the target increment mass m_{inc} of 12 kg, according to Eq. (12), rounding up to time intervals of 0.5 s. This resulted in a t_{inc} value of 4 s.

$$t_{inc\,[s]} = 0.5 \cdot \left\lceil 2 \cdot \frac{m_{inc\,[kg]}}{\dot{V}_{[m^3/s]} \cdot \rho_{[kg/m^3]}} \right\rceil \tag{12}$$

The first sample was taken after an operation time of 5 min, three for averaging mass flow and two for performing the corresponding calculation and for instructing the sampling teams.

Having four sampling teams of two people each and four sampling devices, a sampling interval of 30 s was feasible. The assignment of the increments to the ten samples was alternated, leading to a sampling interval of 5 min for each sample. After a total time of 107 min, the experiment was completed.

Sample processing

Each sample taken was collected in seven waste disposal bins of 220 l each. From there, the path to the particle size–material matrix is a sequence of mass reduction, screening and sorting, as shown in Fig. 3. The single steps are described in the following subchapters.

Mass check and mass reduction

The lower the sample mass, the lower the efforts for screening and sorting. Thus, the primary samples, as well as fine fractions produced through screening, are subject to a mass check. There, it is evaluated whether mass reduction is applicable. Reasons for masses higher than needed are high momentary throughputs during primary sampling time intervals, as well as low coarse fraction shares, while minimum fine fraction masses are recalculated with the new maximum particle diameter, according to Eq. (11). In this work, it was

Fig. 3 Sample processing flowsheet
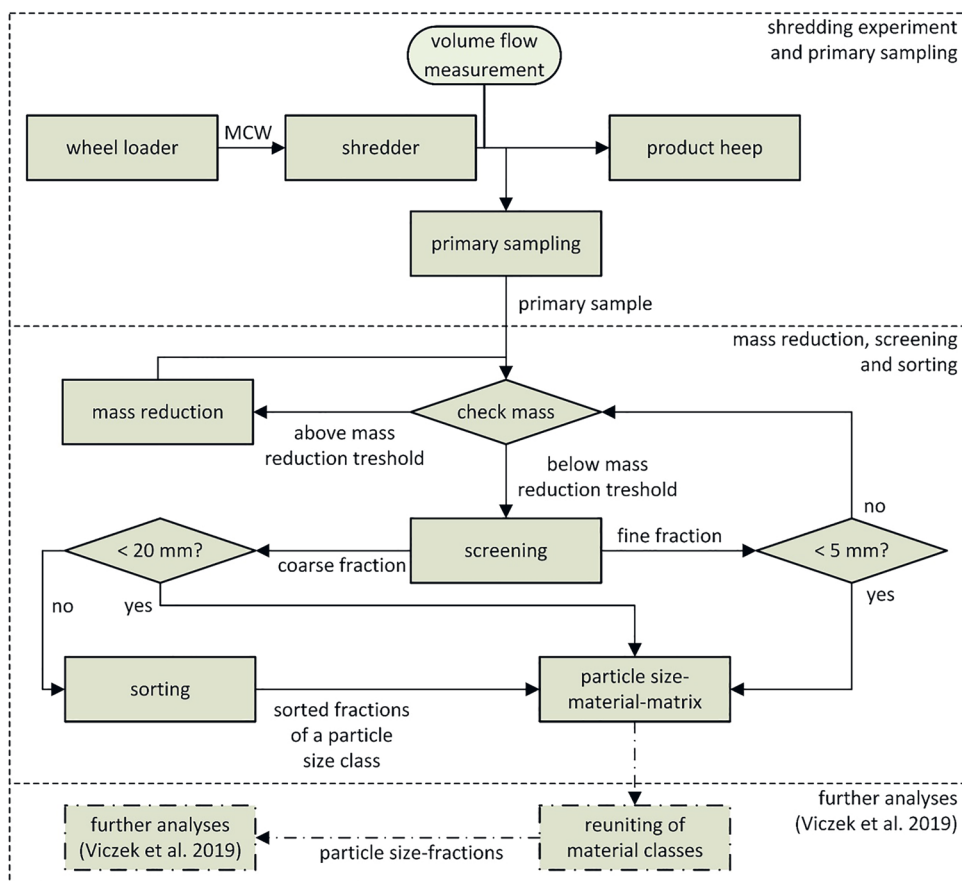
decided to apply mass reduction, when at least 30% of the material can be discarded while adding a safety buffer of 10% of the minimum mass to be kept. This is the case if the inequality shown in Eq. (13) is true for the mass of the evaluated sample part m_{part}, the maximum particle diameter d_{max}, as defined by the preceding screening step (400 mm for the primary sample), and a mass reduction factor f_{red} [–] of 0.7. Discarding less material was considered as not being feasible due to the needed effort for mass reduction. The mass reduction factor defines the fraction of the sample that is preserved. Three possible values for f_{red} were defined, being 0.5, 0.6, and 0.7, respectively. The lowest valid one according to Eq. (13) is chosen. Lower values were not applied, to support spatial coverage of the sample part by taking at least five increments when applying mass reduction, which is a process of subsampling. Being such, it needs to be kept in mind, that all of the described potential sampling errors also apply to subsampling and are added to the primary sampling error. On the other hand, the latter is smaller for these fractions than for coarse ones, due to lower average particle masses, leading to more contained particles per mass.

$$m_{part\,[kg]} \geq \frac{0.66 \cdot d_{max\,[mm]}}{f_{red\,[-]}} \quad (13)$$

Various implementations of mass reduction were described and evaluated by Petersen et al. (2004). Most of them, like riffle splitters, revolver splitters or Boerner dividers are not applicable, as they require the material to be pourable. This is not the case with coarse MCW, which has a high agglomeration tendency. Others, like alternate or fractional shovelling (Petersen et al. 2004) or coning and quartering (Wagner and Esbensen 2012), which are often used for waste mass reduction, show high sampling errors according to the references. Therefore, a mass reduction procedure based on the method of bed blending as described by Wagner and Esbensen (2012) was applied—a comparable approach was used by Pedersen and Jensen (2015) while sampling impregnated wood waste in Denmark.

The sample to be reduced was emptied onto a plastic foil to avoid contamination from the floor or possible loss of material. On this foil, it was spread out, forming an evenly distributed windrow of a length of 5 m (Fig. 4). Using random numbers from 1 to 10 a number of 5–7 segments (0.5 m each)—depending on f_{red}—was chosen for preservation. The other segments were extracted from the windrow and pushed to the floor beside the plastic foil, using a broom. To support this, lines with a distance of 0.5 m were drawn on the

Fig. 4 Windrow for mass reduction

floor next to the foil and on the foil in advance for increment delimitation. The material to be preserved as well as the material to be discarded was shovelled into containers to be weighed. From these masses (i.e. preserved: m_{pres}, discarded: m_{disc}) the real reduction factor f^*_{red} [–] is calculated according to Eq. (14).

$$f^*_{red} = \frac{m_{pres}}{m_{pres} + m_{disc}} \tag{14}$$

Screening

The samples were screened into nine different particle size classes, using screen plates with circular holes of eight different diameters. Screening was performed using a batch drum screen, which has the shape of an equilateral octagonal prism formed by the screen plates. The dimensions are shown in Fig. 5. Screen plates with the following hole diameters were used (in mm): 200, 100, 80, 60, 40, 20, 10, and 5. The screen was operated using material batches of 75 l

for screen cuts of 20–200 mm. The volumetric batch size was reduced by 50% for the smaller ones, as the high mass of the material, caused by the higher bulk density of finer fractions, would have overworked the motor of the screen. Screening times were chosen based on experience, ensuring mass constancy: 180 s for screen cuts of 40–200 mm and 270 s for the smallest three screen cuts. The rotation speed of the screen was set to 5 rpm.

Sorting and mixing

Coarse fractions produced in the screening steps that have particle sizes larger than 20 mm (i.e. in total six fractions), were hand-sorted into nine different material classes. Finer material (i.e. three fractions having particle sizes smaller than 20 mm) was not sorted due to infeasibility, considering the immense amount of work needed to sort such fine materials. Furthermore, finer materials are rarely sorted in practical analyses, as processing them in treatment plants for extracting valuable materials is often disproportionately costly. The material classes were chosen in regard to potential valuables (i.e. from the waste management point of view) contained in the waste, they are: metals (ME), wood (WO), paper (PA), cardboard (CB), plastics 2D (2D), plastics 3D (3D), inert materials including glass (IN), textiles (TX), and a residual fraction (RE). Fractions finer than 20 mm were assigned to the residual fraction. After weighing the sorted fractions, they were joined again, as the subsequent chemical analysis was performed for each particle size class, but not for individual sorting fractions.

Weighing

The weighing was carried out using two different scales. The bigger scale was used for containers with a filled weight higher than 30 kg. With a maximum container weight of about 16 kg, this corresponds to partial samples with a weight of up to 14 kg. The uncertainty of this scale is 100 g. Lighter containers were weighed using a scale with an uncertainty of 0.1 g.

In practice, this means that screening fractions and mass reduction fractions were usually weighed using the big scale. Sorting results were always weighed using the more precise small scale.

Calculations

Particle size–material matrices

The particle size–material matrices regarding masses **M** are $v \times w$ matrices, where the elements m_{ij} represent the mass of
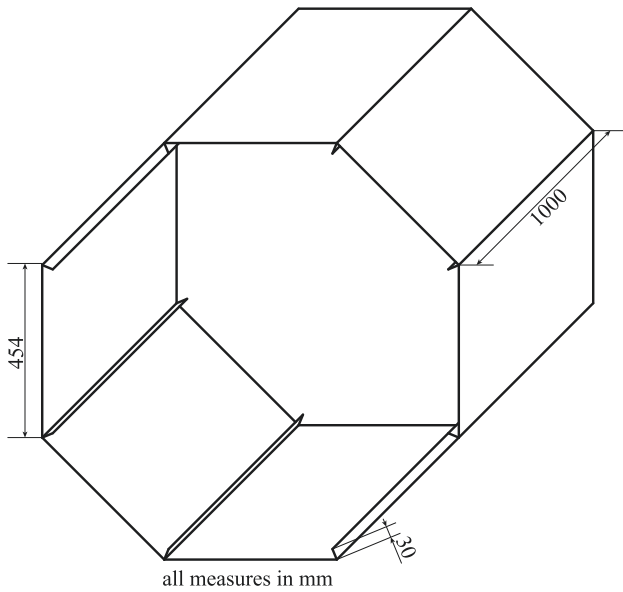
Fig. 5 Dimensions of the screening drum

the ith of v particle size classes and the jth of w material classes in an original sample. The assignment of the indices to the classes is shown in Table 1. For calculating the masses in the original sample, mass reduction steps must be considered mathematically. Thus, the mass m_{ij} is calculated according to Eq. (15), where m_{ij}^* is the mass weighed in the sorting analysis and $f_{\text{red},r}^*$ is the real mass reduction factor (according to Eq. 14), when reducing the fine fraction produced in the $(r - 1)$th screening step. $r = 1$ stands for the original material (0–400 mm) and $f_{\text{red},r}^*$ is 1 if no mass reduction was performed.

$$m_{ij} = \frac{m_{ij}^*}{\prod_{r=1}^{i} f_{\text{red},r}^*} \tag{15}$$

The elements w_{ij} of the particle size–material matrices regarding mass fractions $\mathbf{W}$ represent the shares of masses m_{ij} of the total sample mass and are calculated according to Eq. (16):

$$w_{ij} = \frac{m_{ij}}{\sum_{i=1}^{v} \sum_{j=1}^{w} m_{ij}} \tag{16}$$

As the definition of material classes as well as the choice of screen cuts is arbitrary, further classes can be defined and evaluated by summing up the masses of specific fractions. This allows calculating standard deviations and RSV values for the mass shares of larger fractions, up to the total material. To calculate all different combinations of materials, a binary matrix $\mathbf{B}_w$ containing all possible combinations of ones and zeros for w digits is needed. To generate it, the j column of the matrix contains the w-digit binary representation of the number j, while each of the w rows contains one digit. For w digits, the matrix has $(w^2 - 1)$ columns. For the actual data, w is 9. Equation (17) shows the matrix $\mathbf{B}_4$ as an example.

$$\mathbf{B}_4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{17}$$

Regarding particle size, only adjacent classes are combined, corresponding to an alternating choice of screen cuts. For v particle size classes a binary matrix $\mathbf{A}_v$ is needed, containing all possible combinations of zeros and 1 to v adjacent ones in its rows. The matrix has v columns and $(v \cdot (v + 1)/2)$ rows. For the present data, v is 9. Equation (18) shows the matrix $\mathbf{A}_4$ as an example.

$$\mathbf{A}_4 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}^{\mathrm{T}} \tag{18}$$

The matrices $\mathbf{W}_{\text{mp}}$, containing the weight fractions of all possible combinations of materials and adjacent particle sizes, are calculated according to Eq. (19) and have 45 rows and 511 columns.

$$\mathbf{W}_{\text{mp}} = \mathbf{A}_9 \cdot \mathbf{W} \cdot \mathbf{B}_9 \tag{19}$$

Table 1 Assignment of indices to particle size classes (i) and materials classes (j)

index i/j	1	2	3	4	5	6	7	8	9
Size class [mm]	200–400	100–200	80–100	60–80	40–60	20–40	10–20	5–10	0–5
Material class	ME	WO	PA	CB	2D	3D	IN	TX	RE

Results and discussion

Process and material analysis and primary sampling mass

The mean throughput—determined at the end of the experiment—was 25.2 t/h or 71.8 m^3/h, resulting in a total mass of 45.0 t, a total bulk volume of 128.1 m^3 on the product conveyor belt and thus a bulk density of 351.4 kg/m^3. This means that the bulk density of the shredded material of the REx is much higher than the expected density of 161.8 kg/m^3. Still, the target primary sample masses were approximately achieved, with a mean of 241 kg, a standard deviation of 22 kg, a minimum of 215 kg, and a maximum of 284 kg. The weighted mean values of the particle size–material matrix as well as of the sums of size classes and material classes are shown in Table 2.

Sampling error

The relative sampling variabilities related to the material classes in Table 2 are shown in Table 3. Beyond that, Fig. 6 shows the RSV values as well as the standard deviations s for all 22,995 classes in the matrices $\mathbf{W}_{mp}$, plotted against the correspondent weighted mean values. The grey triangles mark the data corresponding to the original matrices $\mathbf{W}$.

Table 2 Weighted means of particle size–material fractions' mass shares

Particle class [mm]	ME (%)	WO (%)	PA (%)	CB (%)	2D (%)	3D (%)	IN (%)	TX (%)	RE (%)	Sum (%)
0–5	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	17.1[a]	17.1
5–10	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	6.5[a]	6.5
10–20	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	0.0[a]	11.0[a]	11.0
20–40	0.7	1.4	0.8	2.1	0.4	2.3	2.1	0.1	2.2	12.1
40–60	0.8	1.8	0.9	2.5	0.5	2.0	1.0	0.1	2.2	11.8
60–80	1.0	2.0	1.0	2.1	0.7	2.3	0.5	0.5	2.2	12.3
80–100	0.7	1.0	0.7	1.3	0.6	1.7	0.2	0.4	1.4	7.9
100–200	1.0	0.9	0.9	4.7	2.2	3.3	0.4	1.7	2.6	17.6
200–400	0.1	0.0	0.0	0.2	1.1	0.7	0.0	1.1	0.7	3.8
Sum	4.4	7.1	4.3	12.8	5.5	12.2	4.1	3.9	45.8	100.0

[a]Particle size fraction was not sorted: complete material was assigned to the residual fraction with: metals (ME), wood (WO), paper (PA), cardboard (CB), plastics 2D (2D), plastics 3D (3D), inert material (IN), textiles (TX) and residual fraction (RE)

Table 3 RSV values

Particle class [mm]	ME (%)	WO (%)	PA (%)	CB (%)	2D (%)	3D (%)	IN (%)	TX (%)	RE (%)	Sum (%)
0–5	-	-	-	-	-	-	-	-	12.3	12.3
5–10	-	-	-	-	-	-	-	-	12.3	12.3
10–20	-	-	-	-	-	-	-	-	10.4	10.4
20–40	41.4	17.7	24.3	39.3	18.4	17.1	19.7	29.3	22.7	11.6
40–60	47.3	21.5	16.8	25.6	14.2	8.7	37.2	43.4	16.4	8.8
60–80	39.4	23.3	22.9	18.1	17.7	10.0	49.9	30.4	8.9	8.1
80–100	62.0	34.7	38.0	14.2	19.3	17.5	210.7	43.4	17.2	7.7
100–200	74.0	47.7	69.0	21.6	28.9	35.2	131.8	40.9	40.0	10.9
200–400	153.0	230.9	203.9	126.2	38.3	39.8	-	42.9	52.2	28.8
Sum	16.4	18.3	10.5	15.0	16.6	12.1	31.2	26.6	3.6	0.0

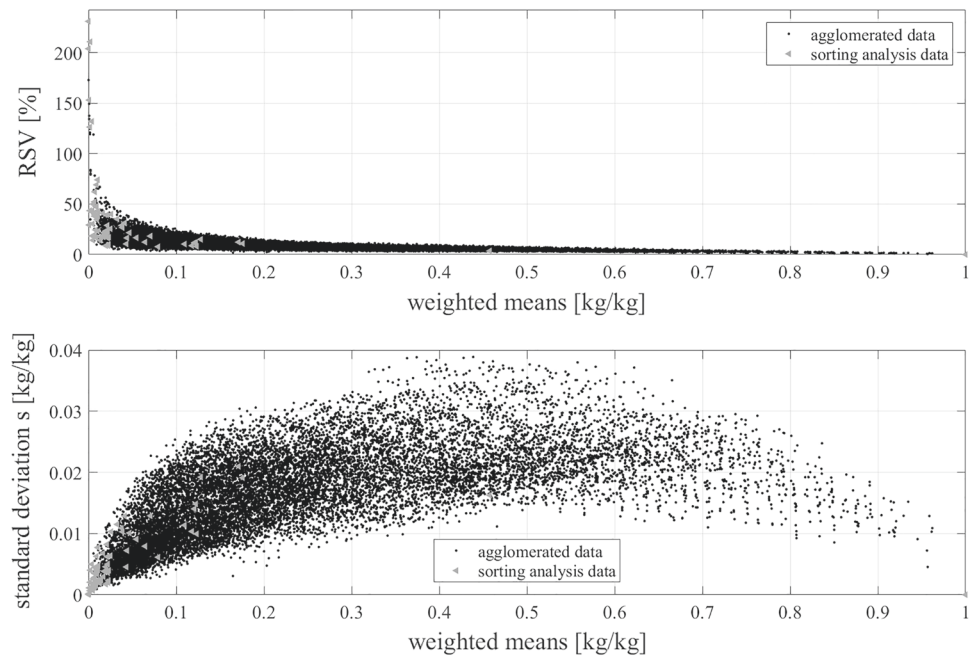RSV < 20%   20% ≤ RSV < 50%   RSV ≥ 50%

Fig. 6 RSV and standard deviation versus weighted mean of mass shares

Incorrect sampling errors

Primary sampling

Increment delimitation is done by defining time intervals during which all material falling from the product conveyor belt of the shredder is collected. Therefore, no IDEs are expected.

Correct increment extraction, on the other hand, turns out to be challenging: some increments could not be taken at the defined time, because the end of the conveyor belt was too high. The reason for this was the wheel loader feeding material into the shredder, not allowing the latter to move forward. Because of this, the belt had to be elevated, producing a higher heap. In these situations, increment extraction started as soon as it was possible again. Furthermore, communication between the samplers and the person responsible for timing was difficult during sampling, due to the loudness of the machine. Because of this, the end of the defined 4 s of sampling was determined by the samplers through counting. Considering the real mass flow of 25.2 t/h, the average sample mass with 20 increments of 4 s each would have been 560 kg, while the observed average was less than half of that, i.e. about 241 kg. Consequently, it can be assumed that the real sampling time was less than 2 s, because of the subjective sense of time, which might also have been influenced by the weight of the increments taken. Nonetheless, this did not negatively affect the target sample mass of 240 kg. Furthermore, all samples fulfil Eq. (11) as the empirical value for d_{95} is 194 mm (calculated through linear interpolation from Table 2). Still, the deviation from the defined sampling time

of 4 s is not very likely to be uniform, leading to scattering of real sampling time and thus to IEE.

Regarding IPE, it cannot be assured that all particles of the taken increments reached the final samples, as handling the bulky sampling device, which is also heavy when filled, might have led to unintentional falling out of some particles.

Mass reduction

Mass reduction is a subsampling process. Consequently, all potential sampling errors might as well occur at this step. Regarding increment delimitation, drawing the equidistant lines on the foil is a quite exact process. So—if IDEs occur at all—the order of magnitude should be negligible compared to IEE:

Extracting the increments correctly turned out to be problematic, especially for coarse fractions. This is because of material wedging, impeding pourability. Because of this, when pushing the segments to be discarded from the foil, it was unavoidable to extract material from the neighbouring segments as well. Therefore, IEE could not be completely avoided with the mass reduction method applied.

Regarding increment preparation, the main expectable error is loss of dust blown away when handling the material.

Screening and sorting

As screening and sorting are not sampling operations, IDE and IEE cannot occur. IPE, on the other hand, are expected

Fig. 7 Relative standard error caused by FSE versus mass share of a constituent c of a two-component composition, according to Eq. (7), for different average particle masses m_c of constituent c with a lot mass of 45,000 kg, a sample mass of 240 kg and an average particle mass for the other constituent k of 0.1 kg
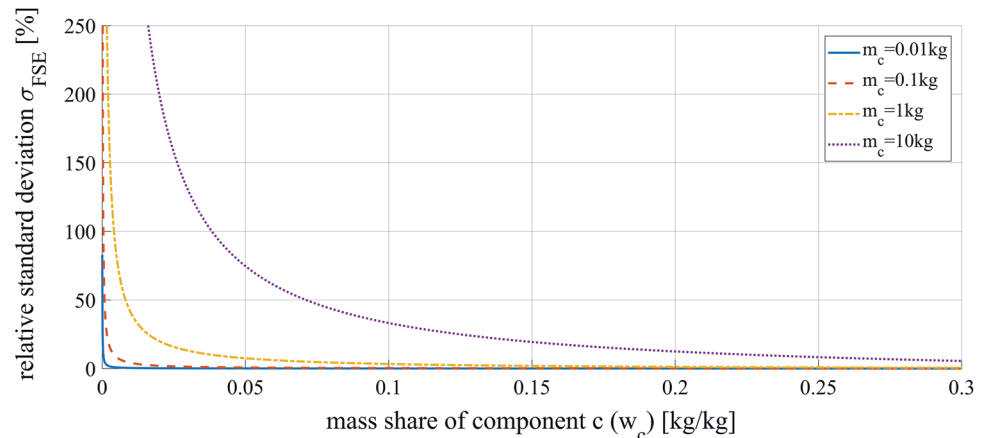
due to blowing away of dust, as well as the loss of particles falling from the sorting table unnoticed.

Loss of water

The samples taken were processed during several weeks. Although stored in closed disposal bins, loss of humidity is possible as the bins are not hermetical, leading to IPE.

The total loss of material, calculated comparing the sum of the individual fraction masses according to Eq. (14) to the primary sample masses, shows a mean value of 5.6% and ranges between 4.6 and 7.2%.

Container and equipment contamination

For practical reasons, before reusing containers, and equipment like the screen or shovels, they could only be cleaned using hand brushes. Thus, cross-contamination of different samples and subsamples cannot be completely excluded, leading to further potential IPE. These contaminations are expected to be low, due to small contact areas in relation to sample masses.

Sampling quality

Applying an RSV of 20% as a threshold for good sampling, Table 3 shows that the applied procedure only produces good results for some of the examined fractions. Figure 6 further shows that RSV tends to be better for large fractions, indicating that small fractions require better sampling and analytics for achieving acceptable relative errors. This is the case, although the absolute standard deviation seems to increase with mass share, reaching a maximum for fraction ratios of 50%.

Figure 7 shows the RSV contribution of the FSE over the mass share (according to Eq. 7) of a constituent c for different average particle masses m_c for the present lot mass and target primary sample mass, assuming a two-component composition. For the average particle mass of the other constituent m_k a value of 0.1 kg was chosen. m_k has little influence on σ_{FSE}, as long as it is significantly smaller than the sample mass m_{sam}. Comparing Figs. 6 and 7, it is apparent that the general trend of the empirical RSV shows similarities to the trend of the FSE's contribution to the sampling error. Especially for coarse particles which are either heavy (e.g. metal 100–200), have very low mass shares (e.g. paper 200–400), or both (e.g. metal 200–400), the FSE explains very well why RSVs far beyond 20% were observed.

However, keeping in mind that only few particle size-material fractions—if any—have average particle masses as high as 1 kg, the figures show that primary sampling FSE only explains part of the observed sampling errors. For example, for a fraction with a mass share of 0.1 σ_{FSE} is 3.4% in Fig. 7, while the RSV values in Fig. 6 range somewhere between 5 and 25%. Therefore, other sampling errors obviously also show significant contributions. Wavrer (2018) highlights the high distributional heterogeneity of municipal solid waste. For MCW, it is known to be even higher, therefore the GSE is likely to significantly contribute to the sampling error. Furthermore, the described ISEs, as well as CSEs and ISEs from subsampling also contribute to the observed RSVs. These contributions, along with the different average particle masses and numbers of subsampling stages for the different data points, as well as errors in estimating RSV due to the low number of 10 taken samples, also explain the scattering of the data in Fig. 6 along the ordinate axis.

Knowing the standard deviation, confidence intervals for estimated concentrations can be calculated as a corresponding

measure. But doing so, care has to be taken for very large RSVs: the high relative errors indicate a positively skewed distribution of the values (and therefore not a normal distribution), as commonly used confidence intervals like 95% would otherwise include negative percentages. Ultimately, the quality of sampling needs to be rated dependent on the analytical target.

Conclusion

A sampling and sample processing procedure for screening and sorting analysis of shredded mixed (i.e. commercial) solid waste was established and is reported, based on the TOS, the Danish horizontal sampling standard DS 3077 and the Austrian standard ÖNORM S 2127. Assessment of sampling quality, rated through the relative sampling variability, shows that the procedure gives good results for some values of the particle size-material matrix (at a threshold of 20%) but not for all of them—it gets as bad as 231%. It is further shown that the RSV is better for larger fractions. In conclusion, especially when analysing small fractions, a reduction in the occurring sampling errors is necessary. Regarding the CSEs, in a first step, it should be assured that the sample mass suffices to keep the FSE within a reasonable range. For this, building a database of typical particle masses, as suggested by Wavrer (2018), is highly encouraged. Equation (7) then provides a (necessary, but not sufficient) minimum sample mass.

As it was shown that FSE only contributes a part of the observed RSVs, compensating the high distributional heterogeneity of the waste by increasing the number of increments might also significantly improve sampling quality by reducing GSE. To handle the resulting higher sampling frequencies, automated primary sampling, e.g. using a reversible conveyor belt, is encouraged.

Such automated sampling might as well contribute to reducing ISEs, i.e. IEEs, as it allows to extract samples exactly in time, while still preserving the benefits of sampling from a falling stream (one-dimensional sampling and good separation of agglomerated particles). Moreover, IEEs during mass reduction (caused by agglomerations, especially for coarse fractions) could as well be reduced by (automatic) subsampling from a falling stream. Furthermore, the analytical error can be decreased by using a more precise big scale.

Finally, when evaluating very small fractions through screening and sorting, increasing sample masses will be unavoidable for reliable analyses. In addition to the discussed estimation of the FSE, the determined values for the RSV help to estimate them in advance.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Austrian Standards Institute (2011) ÖNORM S 2127 Grundlegende Charakterisierung von Abfallhaufen oder von festen Abfällen aus Behältnissen und Transportfahrzeugen [Basic characterization of waste heaps or solid wastes from containers and transport vehicles]. Austrian Standards Institute, Vienna. ICS: 13.030.01

Danish Standards Foundation (2013) DS 3077 Representative sampling—Horizontal standard. Danish Standards Foundation, Charlottenlund. ICS: 03.120.30, 13.080.05

Esbensen KH, Julius LP (2009) Representative sampling, data quality, validation—a necessary trinity in chemometrics. In: Tauler R, Walczak B, Brown SD (eds) Comprehensive chemometrics: chemical and biochemical data analysis, 1st edn. Elservier, Burlington, pp 1–20

Esbensen KH, Wagner C (2014) Theory of sampling (TOS) versus measurement uncertainty (MU)—a call for integration. TrAC. https://doi.org/10.1016/j.trac.2014.02.007

European Committee for Standardization (2006) CEN/TR 15310-1 characterisation of waste—sampling of waste materials—part 1: guidance on selection and application of criteria for sampling under various conditions. European Committee for Standardization, Brussels. ICS: 13.030.10, 13.030.20

Felsenstein K, Spangl B (2017) Richtlinien für die statistische Auswertung von Sortieranalysen und Stückgewichtsanalysen [Guidelines for statistical evaluation of sorting and particle mass analyses], pp 34–68. http://www.argeabfallverband.at/fileadmin/bilder/PDFs/Anhaenge_RA_OOE_2018.pdf. Accessed 19 Mar 2019

Gy P (2004a) Sampling of discrete materials: II. Quantitative approach—sampling of zero-dimensional objects. Chemom Intell Lab Syst 74(1):25–38. https://doi.org/10.1016/j.chemolab.2004.05.015

Gy P (2004b) Sampling of discrete materials—a new introduction to the theory of sampling: I. Qualitative approach. Chemom Intell Lab Syst 74(1):7–24. https://doi.org/10.1016/j.chemolab.2004.05.012

Pedersen PB, Jensen JH (2015) Representative sampling for a full-scale incineration plant test—how to succeed with TOS facing

unavoidable logistical and practical constraints. TOS Forum 3(1):213–217

Petersen L, Dahl KD, Esbensen KH (2004) Representative mass reduction in sampling—a critical survey of techniques and hardware. Chemom Intell Lab Syst 74(1):95–114. https://doi.org/10.1016/j.chemolab.2004.03.020

Viczek SA, Khodier K, Aldrian A, Sarc S (2019) Sampling and analysis of coarsely shredded mixed commercial waste; part II: grain size dependent elemental distribution. In preparation for submission in a peer reviewed process (status Aug 2019)

Wagner C, Esbensen KH (2012) A critical review of sampling standards for solid biofuels—missing contributions from the theory of sampling (TOS). Renew Sustain Energy Rev 16(1):504–517. https://doi.org/10.1016/j.rser.2011.08.016

Wavrer P (2018) Theory of sampling (TOS) applied to characterisation of municipal solid waste (MSW)—a case study from France. TOS Forum 5(1):3–11

Publication II

Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: coarse-shredding of commercial waste

K. Khodier, C. Feyerer, S. Möllnitz, A. Curtis, R. Sarc

Waste Management, 2021, 121, 164–174, https://doi.org/10.1016/j.wasman.2020.12.015

Author contributions according to the CRediT system:

KK: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft preparation, writing – review & editing, visualization, project administration

CF: conceptualization, methodology, writing – original draft preparation, writing – review & editing, project administration

SM: investigation, writing – review & editing

AC: methodology, investigation, project administration

RS: conceptualization, methodology, writing – review & editing, supervision, funding acquisition

Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: Coarse-shredding of commercial waste

K. Khodier [a], C. Feyerer [c], S. Möllnitz [b], A. Curtis [b], R. Sarc [b,*]

[a] Department of Environmental and Energy Process Engineering, Chair of Process Technology and Industrial Environmental Protection, Montanuniversitaet Leoben – Franz-Josef-Straße 18, 8700 Leoben, Austria
[b] Department of Environmental and Energy Process Engineering, Chair of Waste Processing Technology and Waste Management, Montanuniversitaet Leoben – Franz-Josef-Straße 18, 8700 Leoben, Austria
[c] Komptech GmbH, Kühau 37, 8130 Frohnleiten, Austria

ARTICLE INFO

ABSTRACT

Deriving significant experiment-based conclusions on mechanical processing of mixed solid waste is challenging: the input material cannot be downscaled in a way that enables drawing transferable conclusions from lab-scale experiments. Hence experiments need to be conducted in industry-scale, using real waste. Besides the enormous resulting experimental efforts and costs, which economically limit the number of experimental runs, identifying and quantifying significant effects is complicated by the distortion of the data introduced by the waste's variability. The distortion is particularly high for cases where sampling is necessary and in experiments where material cannot be re-used from one run to the next. In the latter case, inter-experimental differences of the waste add to the distortion of the data. In this work, a systematic approach for deriving representative and significant results at the minimum possible effort is described and evaluated, based on the method of Design of Experiments. It is applied to a 32 runs D-optimal industry-scale coarse-shredding experiment with mixed commercial solid waste, based on a reduced cubic design model, examining the influence of the gap width, shaft rotation speed, and cutting tool geometry on the throughput behavior and energy demand. The resulting models are highly significant (model p-values < 0.0001), proving the ability to extract reliable information from industry-scale waste processing experiments. Concerning commercial waste shredding, the models provide new insights into process behavior, for example, the quadratic dependence of the mass flow on the shaft rotation speed, with the highest hourly mass flows at 84% of the maximum shaft rotation speed.

1. Introduction

Mixed solid municipal and commercial wastes represent a significant share of the total waste produced each year, with 13.5 million metric tons (163 kg/inhabitant) alone in Germany in 2018, for example (Umweltbundesamt, 2020). Hence, efficient and effective mechanical treatment processes are needed to enable recycling or energy recovery. This topic's relevance is emphasized by the European Union's recycling targets (2018) of the circular economy package – demanding a recycling rate of 65% for municipal waste by 2035.

For improving the performance of the corresponding treatment processes, it is essential to establish reliable knowledge about the

influences of machine parameters. Such understanding is not only relevant for the classical (stationary) operation of processing plants, but even more for their dynamic operation in industry 4.0-based smart processing plants (see Khodier et al., 2019; Sarc et al., 2019a). For preparation for recycling, but also the production of waste fuel (refuse-derived fuels – RDF), mixed solid waste (MSW) is usually first treated mechanically (e.g. Müller and Bockreis, 2015). Deriving the desired knowledge on the parameter influences of the corresponding machines is challenging: There are many studies in lab-scale (e.g. Luo et al., 2010 on feed moisture's influence on the shredding of municipal solid waste; Kaufeld et al., 2017 on the impact of two-dimensional materials on waste screening), but these use synthetic feed or examine very fine materials and are therefore not directly transferable on real conditions and often even lack statistical analyses of the results. There are also numerous numerical studies (e.g. Sinnott and Cleary, 2015 on

Nomenclature

ANOVA	analysis of variance
c	cutting tool geometry
c_1, c_2	coded representation of the cutting tool geometry
df_{mod}	degrees of freedom of the model
df_{res}	degrees of freedom of the residuals
DFFITS	difference in fits
DFBETAS	difference in coefficients
DMFMS	digital material flow monitoring system
DoE	Design of Experiments
E	mass-specific energy demand
i	number of the actual run
j	factor exponent
k	factor exponent
$K_{jkmn}^{(r)}$	model constant for the factor or interaction $w^j s^k c_1^m c_2^n$ and the response r
m	factor exponent
$\dot{m}$	normalized mean mass flow during an hour of operation
$\dot{m}_{10}$	10th percentile of the momentary mass flow
$\dot{m}_{90}$	90th percentile of the momentary mass flow
n	factor exponent
N	number of runs
p	empirical significance
p_{lof}	empirical significance of the lack-of-fit

p_{mod}	empirical significance of the model
r	arbitrary response variable
R^2	coefficient of determination
R^2_{adj}	adjusted coefficient of determination
s	shaft rotation speed
SS_{mod}	sum of squares of the model deviation from the measured mean
SS_{res}	sum of squares of the residuals
$\dot{V}$	normalized mean volume flow during an hour of operation
$\dot{V}_{10}$	10th percentile of the momentary volume flow
$\dot{V}_{90}$	90th percentile of the momentary volume flow
w	gap width
$\vec{x}$	vector of the factors (i.e. w, s, c_1, c_2)
$y^{(r)}$	measured value of the response r
$\bar{y}^{(r)}$	mean measured value of the response r
$\hat{y}^{(r)}$	model prediction of the response r
$\hat{y}_i^{(r)}$	model prediction of the response r for the factor settings of run i
α	threshold for type I errors

crushing of minerals; Lee et al., 2008 on impact-breakage of concrete waste; Pieper et al., 2016 on optical belt-sorting; Dong et al., 2017 on sieving). But none of the models described there incorporates the variability of mixed solid waste particles, their interactions, and the different behavior of, for example, plastic films in the processing machines yet. The process feed – the waste – cannot be down-scaled and is too variable for current particle-scale models. Hence, real-scale experimental investigations, using real waste, are needed to enable the direct transfer of results on industry-scale processes. The variability of the waste further limits the extent to which results from single experiments can be generalized. When material analyses are involved, the waste's variability influences the experimental results even more due to the necessity of sampling and thereby introduced errors (see Khodier et al., 2020).

The issue of waste variability gets even more challenging when the examined process significantly changes the waste particles' condition: for shredding, for example, the waste cannot be re-used in other experimental runs. Hence, inter-experimental changes in the waste distort the data and the conclusions drawn from them.

Considering the efforts and costs associated with real-scale experiments, the number of conductible experiments is limited in practice. Therefore, efficient experimental methods are needed that ensure the significance of results, despite the limited number of possible experimental runs and the influence of the variable waste input.

The method of Design of Experiments (DoE, e.g. Siebertz et al., 2010; Dean et al., 2017) allows the efficient derivation of empirical multilinear models and consequent findings on the examined processes from experiments. It also involves tests on statistical significance and the calculation of confidence intervals. While the method has been applied on lab-scale waste processing experiments (e.g. Qin et al., 2009 on the electrostatic separation of printed circuit boards; Kazemi et al., 2016 on composting of municipal solid waste), no literature was found showing its application on industry-scale experiments with real waste. Consequently, examining its ability to derive information from such

experiments with acceptable experimental effort is a relevant research question.

Such an evaluation is performed in this work, based on coarse-shredding of mixed commercial waste – to cover all discussed hurdles and contribute insights on the treatment of a waste stream that has hardly been examined in the technical literature yet.

Shredders usually perform the first process step in mechanical treatment plants for MSW (see e.g. Müller and Bockreis, 2015; Pomberger, 2008; Sarc et al., 2016), which is a step of comminution and therefore also of liberation. In addition to the shredder's task in material processing, it is usually the primary dosing device (Feil and Pretz, 2018) and a significant energy consumer. Hence, it significantly influences the subsequent process and the economic efficiency of the waste's mechanical treatment.

Consequently, the energy demand and the throughput were chosen as dependent variables to be examined. Concerning the throughput, it is considered in terms of mass, which is the basis for billing, as well as volume, as machine performance is usually more influenced by the volume flow (drum screens, for example, are limited by the volumetric degree of filling, according to Coskun et al., 2017). Since machine performance is moreover influenced by momentary throughputs, and not just their average (as discussed, e.g., by Feil et al. 2019 and confirmed by the works of Küppers et al. 2020; Curtis et al. 2020), the steadiness of the throughput is also examined.

While the produced material quality – in terms of particle sizes – is also essential, it is not considered for the following reason: particle size distributions are described as continuous functions or as compositions of discrete particle size fractions. Their resulting non-scalar and constrained character requires treatment through sophisticated multivariate statistical methods, such as multiple multivariate linear regression (e.g. Johnson and Wichern, 2007), multivariate analysis of variance (e.g. Hand and Taylor, 1987), and compositional data methods (e.g. Greenacre, 2019), and would burst the scope of this work. For the results' interpretation, some experimental data on this are nevertheless provided where necessary.

Concerning the independent variables – the parameters – the influences of the radial shredder gap width and the shaft rotation speed were chosen for investigation. Also, different cutting tools were used to relate the order of magnitude of observed effects to the cutting tool geometry's influence and demonstrate the handling of nominal (categorical) parameters.

Based on the experimental results of a 32 runs experiment, it is finally targeted to derive significant and trustworthy models on the influence of the described independent variables on the dependent ones. Thereby it is shown whether reliable information can be drawn from such an industry-scale experiment with real mixed commercial waste, applying DoE. The investigation shows that this type of method and analysis is suitable for gaining essential conclusions to develop mechanical waste processing further.

2. Materials and methods

First, the experimental setup is described, including the process flow and the description of the used shredders and measurement equipment. Then, the experimental design is presented, explaining the choice of the number of runs and the design. Finally, the methods applied for evaluating the results are described.

2.1. Experimental setup

2.1.1. Material

The feed material used in the experiments was waste declared as solid mixed commercial waste from Styria in Austria in October 2019 (photo: Fig. S1; composition: Fig. S2 in the supplementary material). Analyses of similar waste have been published recently: e.g., Möllnitz et al. (2020) with a focus on plastic types; analyses on a particle level by Weissenbach and Sarc (2020), and the sieving and sorting analyses and physical-chemical analyses of RDF, produced in plants which process mixed commercial waste, by Sarc et al. (2019b).

Due to the substantial heterogeneity of this kind of waste (shown e.g. by Khodier et al., 2020), it was attempted to keep its similarity high to decrease the distortion of results, e.g., by limiting the number of collection sources.

2.1.2. Flow chart

Fig. 1 shows the experimental setup. The waste was directly fed into the shredder using a wheel loader. To minimize other machine influences, caused, for example, by the strength of the motor or the geometry of the feeding bunker of the shredder, the same basic type of shredder was used in all cases: a Terminator 5000 SD, which is a single-shaft shredder from the Austrian company Komptech GmbH, that is a research partner within the funded project ReWaste4.0. From the shredder's output belt, the waste was passed to a digital material flow monitoring system (DMFMS), which measures momentary throughput in terms of volume and mass. The waste leaving the device was collected on a product heap.

There are different possible reasons for the potential occurrence of momentary zero-throughputs throughout the experiment. While reversing the shaft is an inherent characteristic of the process, prior experiments showed that the wheel loader sometimes cannot keep up with the shredder at higher throughputs. As these experiments intend to evaluate the shredder parametrization's influence without considering feeding conditions, a camera that shows the inside of the shredder's feeding bunker was installed. The recordings were synchronized with the throughput measurements to evaluate the reasons for periods with zero-throughputs and discard them if they are caused by the absence of material, as proposed by Aarne Vesilind et al. (1980).

2.1.3. Cutting tool geometry

Three different cutting tool geometries of the Terminator 5000 SD were used in the experiments, called "F", "XXF", and "V" (see Fig. 2), demonstrating the consideration of nominal factors. The geometry is described through a rotating shaft system with the primary cutting tools and a counter comb, which functions as the counter cutting edge. The examined systems differ regarding the geometry and the number of cutting tools, and the axial and radial cutting gap width. The latter can be adjusted continuously. The technical data of the three geometries are summarized in Table 1.

The F-type cutting tool has 32 cutting teeth on the shaft, arranged as a double helix. The counter comb contains 17 teeth, which are all positioned at the same height. The cutting process takes place only on the left side of the teeth, which is the reason for the different left-side and right-side cutting gaps.

Unlike the F-type, the XXF-type cutting tool has only 22 teeth on the shaft, arranged as a V-pattern, also called "chevron pattern". The counter comb contains 23 teeth, which are all positioned at the same height as well. Same as with the F-type, the process of cutting takes place only at the left side, but with a much smaller axial gap width.
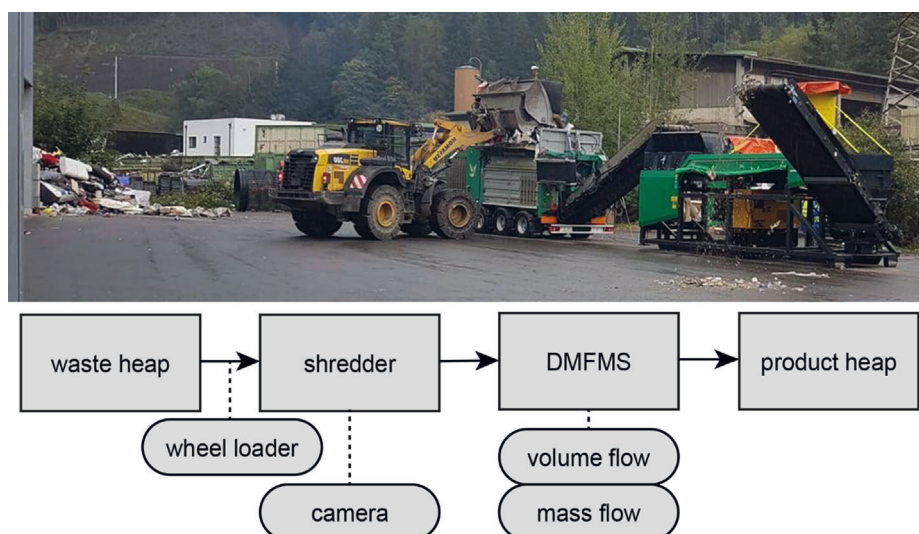

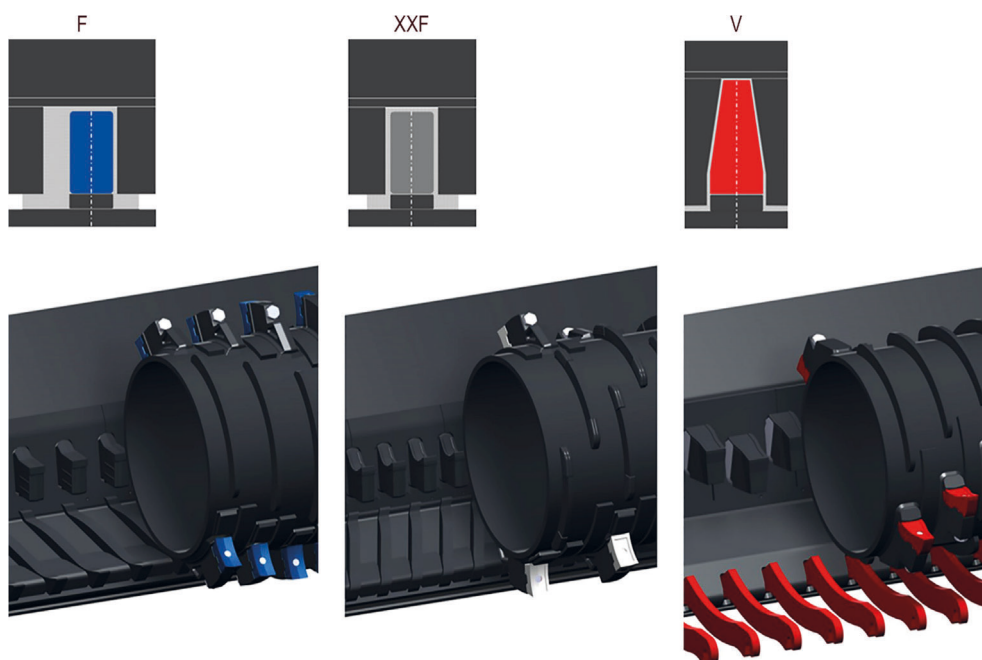**Fig. 1.** Experimental setup: photo and flow chart.

Fig. 2. Cutting tool geometries.

Table 1
Technical data of the cutting tool geometries.

type	F	XXF	V
number of cutting teeth (shaft) [pcs.]	32	22	32
position of cutting teeth (shaft) [–]	double helix	chevron	chevron
width of cutting teeth (shaft) [mm]	70	70	42/85*
height of cutting teeth (shaft) [mm]	124	124	183
width of cutting teeth (counter comb) [mm]	64	54	81/100*
height of cutting teeth (counter comb) [mm]	142	136	202
cutting circle [mm]	1070	1070	1170
length of shredding-shaft [mm]	3000		
right side cutting gap (axial) [mm]	3.5	2	3
left side cutting gap (axial) [mm]	39	2	3
minimum cutting gap (radial) [mm]	0		
maximum cutting gap (radial) [mm]	33	35	30/38*
comb-system [–]	no	no	yes

* bottom/top of the teeth.

The V-type has 32 cutting teeth on the shaft that are arranged as a chevron-pattern as well. As shown in Fig. 2, the shape of the teeth is different from the other two types. Also different is that cutting takes place at both sides of the teeth, resulting in a smaller number of 17 counter comb teeth. These are positioned at alternating heights to facilitate large particles' comminution through the higher bending tensions that are achieved at the larger gap between two of the upper teeth. Additionally, the V-type has a comb-system (see Fig. 2), which supports the production of defined particle sizes and leads to some secondary comminution.

2.1.4. Digital material flow monitoring system (DMFMS)

The throughput is measured using a DMFMS. It incorporates two different measurement systems: the mass flow is measured through an integrated belt scale. The volume flow is determined using a laser triangulation measurement bar above the belt.

The belt scale delivers updated values every 3 s (with an accuracy of ±2% in the range of 25–100% for throughput rates of 5–100 t/h). These values represent the arithmetic average of the mass flow during this period. The laser triangulation system measures the material stream's height profile with a frequency of 200 Hz, delivering updated values every 2 s, representing the arithmetic mean of the volume flow during this time.

2.2. Experimental design

The experiments are performed based on DoE. It aims at efficiently performing experiments to maximize the gained information per experimental run. Unlike traditional experimentation, instead of varying one factor at a time while keeping all others constant, factors are changed simultaneously, enabling a more robust identification of factor influences and the identification of factor interactions (Siebertz et al., 2010). The design is chosen, and the results are interpreted based on an empirical regression model – usually a polynomial multilinear model.

Dean et al. (2017) describe some fundamental principles of DoE: Replication of experimental runs is essential to determine the random variation between single runs, to differentiate between significant effects and random noise, caused, for example, by the differences of the waste input from one experiment to the other. Blocking means to divide the experimental runs into groups of more similar conditions, then achievable for the overall investigation. In terms of waste processing, these groups could be defined by different wheel loader drivers or waste collection sources, for example. Furthermore, randomization is an essential concept in DoE: experiments are preferably performed in random order – so after designing the experiment, the runs' order is shuffled. Thereby, the random spreading of systematic changes over time – e.g., wear of the shredder's cutting tool – over the examined factor ranges and the random spreading of disturbances like waste composition is incorporated.

In the terminology of DoE, according to Siebertz et al. (2010), properties that influence the experimental system (i.e. the shredder) are called parameters (not to confuse with parameters in terms of process parametrization). Some of these can be set on purpose, while others (e.g. the random waste composition) cannot. Those parameters that are chosen to be deliberately set during

the experiment are called factors. The outputs of the system, like the mass throughput of the shredder, are called responses. A change of the magnitude of a factor's effect on a response, with different settings of one or more other factors, is called an interaction and is described through these factors' products in the regression model.

2.2.1. Duration and number of experimental runs

The choice of the duration of each experiment is based on a reasonable compromise between the following: maximizing the inter-experimental homogeneity of the used waste by maximizing the experimental durations and minimizing the experimental duration to increase the number of experiments that can be feasibly conducted, to allow more complex regression models and increase the statistical robustness of the results – in particular, in consideration of the distortion of the data caused by the waste's heterogeneity.

In this study, the maximum practicable number of experiments was also influenced by the effort for material analyses for each experimental run, as material sampling and sieving and sorting analyses were also performed. Ultimately, balancing these arguments, an experimental duration of one hour was chosen while allowing 32 experimental runs.

2.2.2. Factor types and ranges

The experiments investigate the influence of three factors: gap width (w), shaft rotation speed (s), and cutting tool geometry (c). The gap width is a numeric factor. It was described in percent of the maximum aperture, setting the lower limit to 0% and the upper limit to 100%. The gap width is set by hand using an analog display, which does not allow high accuracy. Hence, the factor was defined as discrete, only allowing steps of 10%, resulting in 11 different factor levels.

The shaft rotation speed is a numeric factor as well. It was described in percent of the maximum rotation frequency, which is 31 rpm. 100% of this maximum frequency was defined as the upper limit, while 60% was chosen as the lower limit, based on the manufacturer's recommendations. The software of the machines only allows setting the shaft rotation speed in steps of 10%. Therefore, the factor was also defined as a discrete factor, resulting in 5 possible factor levels.

The cutting tool geometry is a nominal factor. As three types of geometries were examined, it is a nominal factor with three levels more precisely.

In DoE, factors are coded according to Siebertz et al. (2010), using, for example, a range from -1 to 1 for numeric factors. Therefore, a gap width of 0% corresponds to -1, while a gap width of 100% corresponds to 1. For the shaft rotation speed, -1 stands for 60%, while 1 equals a speed of 100%. Nominal factors are coded using one of several versions of so-called contrast matrices. In this work, the cutting tool geometry is represented by two variables c_1 and c_2, based on sum contrasts (see e.g. Chambers and Hastie, 1993). Consequently, the F-type cutting unit is represented by $(c_1, c_2) = (1, 0)$. For the XXF unit $(c_1, c_2) = (-1, -1)$, while for the V-type cutting tool, $(c_1, c_2) = (0, 1)$. The sum contrasts result in mean values of 0 for c_1 and c_2, which is desirable for the analysis of the results.

2.2.3. Design model

The regression model used to design the experiment is a reduced cubic design, according to Equation (1), where $\hat{y}^{(r)}\left(\vec{x}\right)$ is the model value of an arbitrary response r (e.g. the specific energy demand), $K_{jkmn}^{(r)}$ is the model constant for the coded factor or interaction $w^j s^k c_1^m c_2^n$ and the considered response r, and $\vec{x}$ is a vector of

all factors. The model corresponds to the full quadratic model for every factor level of the nominal factor c and therefore for every cutting tool geometry.

$$\hat{y}^{(r)}\left(\vec{x}\right) = \sum_{j=0}^{2} \sum_{k=0}^{2-j} \sum_{m=0}^{1} \sum_{n=0}^{1-m} \left(K_{jkmn}^{(r)} w^j s^k c_1^m c_2^n\right) \tag{1}$$

The choice was made based on the following considerations: Nominal factors can only appear linearly in the regression model, as no factor levels between them exist. Therefore, a more elaborate definition of a curve shape between factor levels does not make any sense. Hence the exponents m and n of the factors c_1 and c_2 do not exceed the value 1. The linear nominal factor terms can still appear as part of higher-order factor interaction terms.

The order of the numeric factors w and s is not limited mathematically. Additionally, it is unknown whether the influence of these factors is linear. Therefore, including higher-order terms might make sense. Conversely, considering the physics of the investigated process, it was assumed to be likely for the considered responses' correlations with numeric factors to be strictly monotonic – no turning points were expected. However, single ones would be covered by the model as well. To avoid an unfeasible extent of the experiment by using the full cubic model, it was further necessary to assume that no inflection points appear. A quadratic model should then suffice to approximate the data well. These hypotheses are later validated through a test for lack-of-fit.

2.2.4. Design and parametrization

There are many standard designs for performing experiments. Examples are Fractional Factorial or Central Composite (Dean et al., 2017). But none of the standard designs is well suited for the specific case described, mainly for the following reasons: The design must consider the limitation that only discrete factor levels can be chosen. Moreover, the chosen model (reduced cubic) is not a standard model. The design should consider it without increasing the number of runs to the level needed to fit, for example, a full cubic model.

In conclusion, an optimal design was chosen for this experiment. This class of designs iteratively optimizes one of various statistical criteria while considering constraints of the design space and arbitrary design models.

To generate the design, the software Design Expert® 11 was used. It offers three of many available different optimization criteria: I-, D- or A-optimal designs. According to the documentation of Design Expert® (Stat-Ease, 2018), as well as Siebertz et al. (2010), D-optimal designs perform best at estimating the effects of factors. They minimize the volume of the coefficients' confidence ellipsoid and therefore give the most robust factor estimations. Consequently, a D-optimal design was used since determining factor strengths is the aim of the experiment.

For generating a D-optimal design, some more parameters must be chosen: the candidate points define the points that the algorithm may choose from when generating the model. Having defined discrete levels for all factors, all points are regarded as categoric ones, resulting in 165 candidate points for the model. Furthermore, the number of blocks must be specified. Potential blocking parameters for this experiment are the batches of waste delivered by every truck and the wheel loader driver feeding the shredder – two drivers supported the experiments. As these parameters' assignment to individual runs was not projectable, no blocking was used, so the number of blocks was set to 1.

The number of replicate points was set to 5. These are used to calculate the pure error. The number of lack-of-fit points was also set to 5. They are used to perform lack-of-fit tests, which evaluate if the chosen kind of model can sufficiently approximate the experimental data. These default settings ensure a powerful test

for lack-of-fit. Four additional model points were used, other than the 18 required model points, to exploit the allowed number of runs since prior experiments indicate that a high number of runs is needed to obtain significant results when performing this kind of experiment with this kind of material. Therefore, the additional model points are expected to increase the reliability of the results significantly.

2.3. Analysis of the results

Analysis and modeling of the results are performed using Design Expert® 11, completing the steps described in the following:

2.3.1. Model reduction

The factors and interactions shown in Equation (1) were chosen before the experiment, aiming at providing a comprehensive mathematical regression model to describe these future observations. Consequently, it was not known at that moment which of these terms significantly influence the responses to be investigated. Hence, model reduction is applied to eliminate non-significant factors, simplifying the model, and facilitating interpretation. It also increases the significance of conclusions drawn, as more degrees of freedom (generated through experimental runs) can be assigned to the residual errors.

In this work, forward selection based on the empirical significance (p-values), choosing a threshold of $\alpha = 0.1$ for type I errors (see Wasserman, 2013) is applied. The algorithm starts from the most basic model (Equation (2), where $\bar{y}^{(r)}$ is the arithmetic mean of the observed responses $y_i^{(r)}$). It then keeps adding terms one after the other, always choosing the most significant design model term that was not included yet, as long as terms with the empirical significance $p < \alpha$ remain. According to Siebertz et al. (2010), it is also essential to ensure model hierarchy. Therefore, after applying forward selection, any lower-order terms included in selected higher-order terms are also included.

$$\widehat{y}^{(r)} = K_{0000}^{(r)} = \bar{y}^{(r)} \tag{2}$$

2.3.2. Analysis of variance

The analysis of variance (ANOVA, see e.g. Siebertz et al., 2010) shows some measures used to evaluate the model. The following of these are reported in this work:

The coefficient of determination R^2 determines how much of the variation of the data around the mean is explained by the model. It is calculated according to Equation (3). SS_{res} is the sum of squares of the residual errors according to Equation (4), where $y_i^{(r)}$ is the observed response r at the ith run, $\widehat{y}_i^{(r)}$ is the predicted response for the factor settings of the ith run, and N is the number of runs. SS_{mod} (Equation (5)) is the sum of squares of the deviations of the model's predictions from the mean response (Siebertz et al., 2010).

$$R^2 = 1 - \left(\frac{SS_{res}}{SS_{res} + SS_{mod}} \right) \tag{3}$$

$$SS_{res} = \sum_{i=1}^{N} \left(\widehat{y}_i^{(r)} - y_i^{(r)} \right)^2 \tag{4}$$

$$SS_{mod} = \sum_{i=1}^{N} \left(\widehat{y}_i^{(r)} - \bar{y}^{(r)} \right)^2 \tag{5}$$

The adjusted coefficient of determination (R_{adj}^2) is also a measure of the share of the data's variation around the mean value explained by the model. But it is adjusted by the number of terms in the model so that it decreases compared to R^2 with an increasing number of terms. Consequently, it is a measure of the model's efficiency in terms of explained variance per term. It is calculated according to Equation (6), where df_{mod} is the number of degrees of freedom consumed by the model (which is equal to the number of terms in the model) and df_{res} is the number of degrees of freedom of the residuals (which is equal to $N - df_{mod}$) (Siebertz et al., 2010).

$$R_{adj}^2 = 1 - \left[\frac{SS_{res}/df_{res}}{(SS_{res} + SS_{mod})/(df_{res} + df_{mod})} \right] \tag{6}$$

The p-value of the model p_{mod} describes the probability of observing the effects that the model describes due to noise rather than real effects. The lower it is, the more significant is the model.

The p-value of the lack-of-fit p_{lof} is desired to be non-significant and therefore larger is better. It describes the probability of observing the present lack-of-fit due to noise.

2.3.3. Model diagnostics

Siebertz et al. (2010) describe different plots that are used to evaluate models. The "Full normal plot" is used to examine the distribution of the residuals, which are the differences between measured responses and the corresponding model values. Their distribution is relevant because the Fisher F-test used in the ANOVA for calculating the p-values requires normally distributed residuals.

There are three more plots, which are used for examining the model predictions according to Siebertz et al. (2010): "Predicted vs. Actual", "Residual vs. Predicted", and "Residual vs. Run". They are used to visually examine the data for systematic trends of the residuals (which are not desired) concerning the predicted value, the run number, and each factor's values. Another relevant plot to examine is the "Residual vs. Factor" plot, which supports identifying systematic relations between the factor settings and the prediction errors.

Furthermore, three measures for the influence of single runs are used for outlier analysis: Cook's distance, DFFITS (difference in fits), and DFBETAS (difference in coefficients, which are often called betas). They calculate changes of the model when not considering a certain point for its calibration: Cook's distance is a measure for changes of the predictions for all design factor settings. DFFITS is a measure for changes in the predictions of the very points that are not considered. DFBETAS are calculated separately for each model coefficient and show changes of its estimate.

After defining the model through model reduction and performing the coefficients' calibration, all these plots are examined. If any of them shows undesired behavior, it is further analyzed and treated, e.g., through transformations of the response variables (for example, logit-transformation or power transformation) or outlier investigation and possible elimination. If any such treatment appears to be necessary, the model finding starts again from model reduction.

3. Results and discussion

3.1. Experimental design and results

The experimental design created by Design Expert® is shown in Table S1 in the supplementary material. Furthermore, it presents the measured mean of the volume flow ($\dot{V}$) and the mass flow ($\dot{m}$) over the experimental duration of one hour, the ratio of the 10th and 90th percentile of both ($\dot{V}_{10}/\dot{V}_{90}$ and $\dot{m}_{10}/\dot{m}_{90}$), and the mass-specific energy demand (E). For reasons of confidentiality, the values of $\dot{V}$ and $\dot{m}$ in this work were normalized (division by

the arithmetic average observation of all runs). Furthermore, there are two columns, "run plan" and "run actual". The first of the two shows the run order, which resulted from the randomization performed by Design Expert®. Throughout the experiments, it was found out that the V-shredder's motor rotation speed had been unintentionally changed from its standard configuration before the planned run number 16, which affected the planned runs 16, 18, and 19. Hence, these runs were dismissed and repeated after discovering the issue. Due to the tight timescale and the time consumed when switching shredders, it was unfortunately impossible to re-randomize the remaining runs, including these three. Consequently, the runs were repeated right after the planned run 24, accepting the resulting impairment of randomness caused by this.

3.2. Modeling of the responses

Table 2 shows the resulting models and coefficients and the reported ANOVA results for all examined responses.

The p_{mod}-values show that highly significant models were found for all responses, with probabilities of less than 0.01% of observing such experimental results due to random noise. The models also show non-significant lack-of-fits at a standard threshold of 0.05 (see e.g. Siebertz et al. 2010), although it is close to critical for the mean volume flow, with a value of 0.0507. The R^2-values are all in a similar range, with values ranging from 0.73 to 0.87. Considering the high expected residual errors caused by the highly variable nature of the waste (see e.g., throughput fluctuations described by Curtis et al., 2020; intra-experimental waste variability described by Khodier et al., 2020), these values indicate that the models explain the observed results quite well.

For E, a significant model with non-significant lack-of-fit was found, containing only the intercept and the cutting tool, but showing a lower R^2 value of 0.6938 compared to the other models and some deviation from normality for the residuals. The model could be improved by introducing a transformation, applying the natural logarithm to the specific energy demand. The new model, which is presented in Table 2, shows the desired normality of the distribution of the residuals and a larger share of observed variance explained by the model. As the displayed R^2 value for the $\ln(E)$-model is calculated, considering $\ln(E)$ as a response, the R^2 for the response E, based on the resulting exponential model was also calculated, for comparison with the R^2 obtained before the

transformation. It has a value of 0.75 and therefore confirms the claimed model improvement.

3.3. Factor influences

Fig. 3 shows the factors' effects (and their corresponding confidence bands) on the modeled responses, setting the two other factors to the average level (coded 0) for every subplot. The models for the mean mass flow and the specific energy demand contain no interaction terms. Changing the corresponding other two factors from the average level consequently only changes the effect plots' absolute location, but not their slope. For the other three responses, the effect of the interactions is shown in Fig. 4. In the following sections, the response values at $w, s, c_1, c_2 = 0$ are called "intercept-value", while the responses for average gap width and shaft rotation speed settings ($w, s, = 0$) for the different cutting tools are denoted as "F-intercept" ($c_1 = 1, c_2 = 0$), "V-intercept" ($c_1 = 0, c_2 = 1$) and "XXF-intercept" ($c_1 = -1, c_2 = -1$).

3.3.1. Gap width

As shown in Fig. 3, a linear influence of the gap width on the mean volume flow was observed. It increases by about 24% of the intercept-value when increasing the gap width from 0% to 100%. The main reason for this increase is presumably the resulting higher available flow area for the waste, which is the major influence on crushers' throughput, according to Cleary and Morrison (2021). Furthermore, the larger gap increases the maximum dimensions of particles that may pass through without being comminuted. Though, the model shows an interaction of the gap width and the cutting tool geometry so that the effect strength can and must be evaluated separately for the three cutting tool types. As shown in Fig. 4, the effect is very small, far from significant, for the V-type cutting tool. In addition to the unlikelihood of a decreasing throughput at increasing gap width, considering domain knowledge, the confidence bands let assume that there is hardly any effect of the gap width for the V-type unit.

On the contrary, for the F and XXF cutting tools, much more significant and similar influences of the gap width were observed. They show an expected increase of 28% of the F-intercept for the F-shredder and 42% of the XXF-intercept for the XXF-type cutting tool. The V-type geometry's different behavior may be explained by the lower increase of the available flow area at larger gaps, due to the conical shape of its teeth, compared to the other geometries. Moreover, the effect of the gap width is most likely

Table 2
Model coefficients and ANOVA results.

| coded model coefficients					
Factor	$\dot{V}[-]$	$\dot{m}[-]$	$\dot{V}_{10}/\dot{V}_{90}[-]$	$\dot{m}_{10}/\dot{m}_{90}[-]$	$\ln(E[L/t])$
1	0.999	1.290	0.270	0.281	0.701
w	0.120	0.112	−0.009	−0.004	−0.111
S	0.135	0.151	−0.039	−0.037	−0.007
c_1	0.274	0.090	−0.074	−0.043	−0.210
c_2	−0.324	−0.215	0.141	0.080	0.420
wc_1	0.056	–	0.013	0.009	–
wc_2	−0.159	–	−0.011	−0.004	–
w^2	–	–	0.014	−0.008	–
s^2	–	−0.353	–	–	0.193
$w^2 c_1$	–	–	0.023	−0.012	–
$w^2 c_2$	–	–	0.069	0.095	–
ANOVA and diagnostics results					
R^2	0.79	0.73	0.87	0.84	0.76
R^2_{adj}	0.74	0.6723	0.81	0.77	0.72
p_{mod}	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
p_{lof}	0.0507	0.6254	0.4173	0.8102	0.3095
removed outliers (Run actual)	–	4, 7	21	–	4
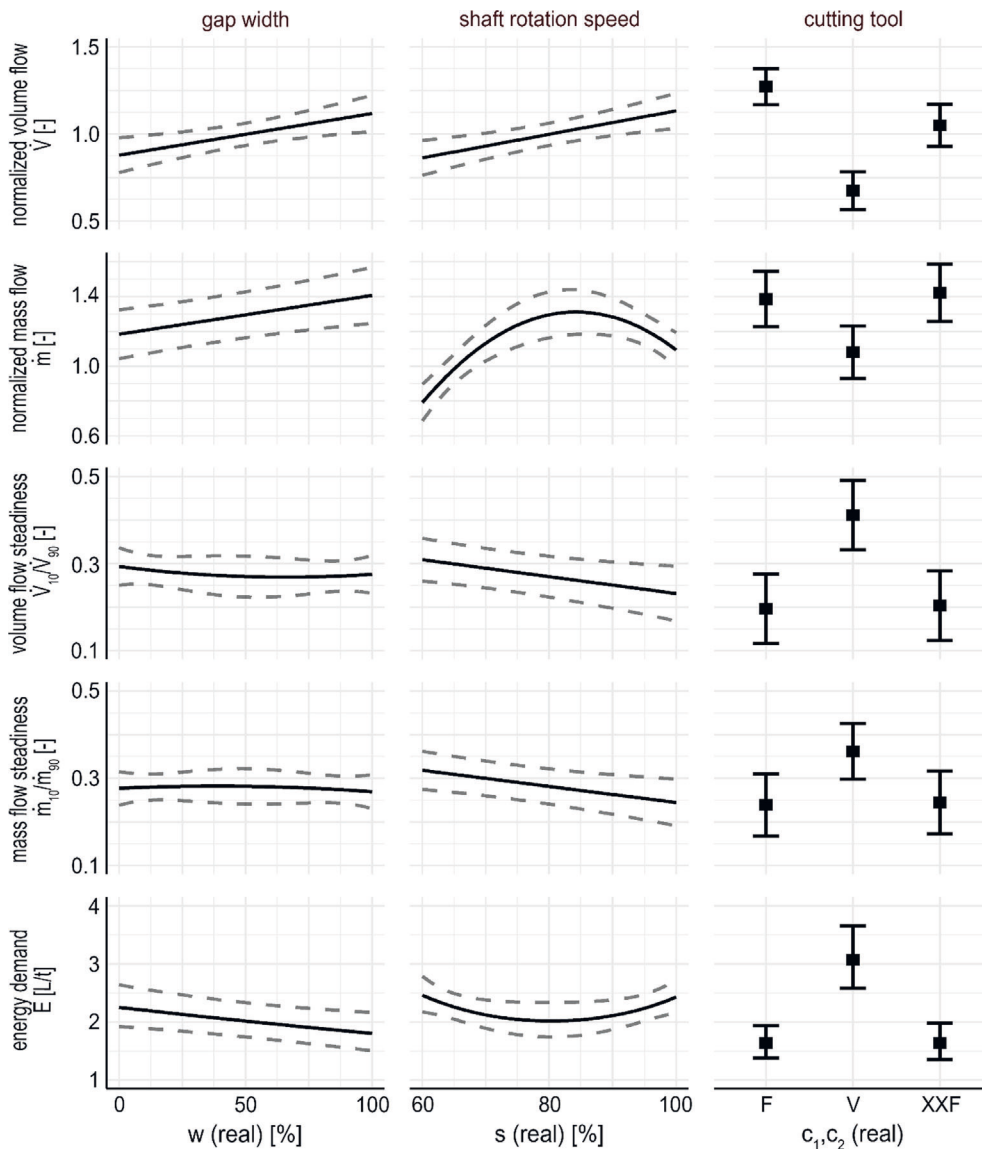
Fig. 3. Effect plots with confidence bands for the influence of gap width, shaft rotation speed, and cutting tool geometry on throughput, throughput steadiness, and specific energy demand, at the average setting of the respective other factors.

significantly weakened by the comb system of this unit, in analogy to the influence of the grid size influence on the throughput of a battery cell crusher, that was reported by Wuschke et al. (2019).

A linear influence was found regarding the mean mass flow, but no significant interactions were observed. Considering the proportionality of volume and mass in average (linked through the average bulk density), it is reasonable to assume that there is also an interaction with the cutting tool geometry for the mass flow, which was not detected due to higher noise in the mass flow data. Consequently, the much wider confidence bands for the influence of the gap width on the mass flow, compared to its impact on the volume flow, result from a combination of higher noise, potential effects on material density, and variation caused by the geometry, that is not explained by the model. Finally, according to the model, the expected value for the increase of the mass flow, when changing the gap width from minimum to maximum, is 17% of the intercept-value.

The models for the influences on throughput steadiness in terms of volume and mass both show quadratic effects with cutting tool geometry-interactions for the gap width. The effect, on average of the three cutting tools, is not very significant, according to Fig. 3. Fig. 4, on the other hand, shows more significant influences, particularly for the V and XXF cutting tools.

Remarkably, a turning point was observed, close to the average gap width of 50%, with increasing values for smaller and bigger gaps for the V-type and decreasing values for the F- and XXF-type shredder. For the F- and XXF-types, a possible interpretation is that for small gap widths, it is harder for the shredders to pull some materials through, leading to fluctuations. In contrast, for large gaps, the ease of passing through the gap might lead to instantaneous passing of large amounts of material from time to time. Still, an in-depth analysis is subject to further research, especially concerning the V-type shredder's behavior, where no explanation for the observed trend was found yet.

As for the specific energy demand, the shape of the curve, resulting from the logarithmic model, shows an almost linear expected decrease (without any interactions) of 0.45 Lt^{-1} when changing the gap width from minimum to maximum. The main explanation for this is the attribution of the idle energy demand per time to a larger amount of waste because of the increased
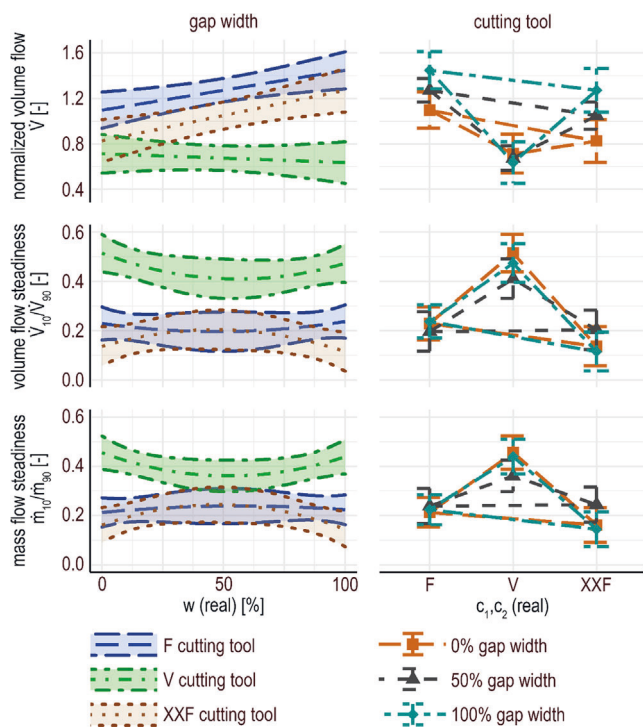
Fig. 4. Interaction plots for the influence of the factors gap width and cutting tool geometry on the mean volume flow and the volume-related and mass-related throughput at different settings of the other factor.

mass-throughput. Furthermore, a potentially increasing share of particles that pass through without comminution, as discussed above, may also contribute. The energy demand for breaking particles, on the contrary, is not expected to change, as it mainly depends on the axial distance between the teeth of the counter comb, according to Feyerer (2020).

3.3.2. Shaft rotation speed

For the shaft rotation speed, no significant interactions were observed. Consequently, the following observations are all related to average effects.

When increasing the shaft rotation speed from 60% to 100%, a linear increase of the volumetric throughput was observed, with an expected value of 14% of the intercept-value.

Regarding the mean mass flow, a quadratic influence of the shaft rotation speed was observed, showing the lowest throughput for the minimum chosen shaft rotation speed of 60%, increasing up to a rotation speed of 84%, and then decreasing again with increasing rotation frequency. A reason could be the material intake behavior at high rotation speeds. The difference between the maximum and minimum expected value is about 40% of the intercept-value, and therefore comparatively high.

In addition to the observed curve's quadratic nature, especially the difference in its trend compared to the mean volume flow raises questions in terms of interpretation. It was examined whether the difference results from the different drawn conclusions during outlier detection, as shown in Table 2. This consideration showed a partial explanation but did not fully explain the observed differences. Removing runs 4 and 7 when modeling the volume flow leads to a significant but weaker quadratic term in the model. On the contrary, considering them in the mass flow model does not remove the quadrature but weaken it. Another interpretation is that the shaft rotation speed might influence the mean density of the product material. But an intent to model such an effect from density data, calculated from the mass and volume

flows in Table S1, failed to show an explanatory significance. Eventually, the potentially critical p_{lof}-value of the volume flow model might indicate, that the mass flow curve is more trustworthy, while more data is needed to validate the trend of the volume flow curve. After all, a confident interpretation is subject to further research.

Regarding throughput steadiness, a negative linear effect of increasing the shaft rotation speed from 60% to 100% was observed for the volume flow (29% of the intercept-value) and the mass flow (26% of the intercept-value). In general, the two curves are very similar, qualitatively, and regarding the absolute values, despite the slightly different update intervals of the sensors, which lead to different smoothing of the curves. The observations are reasonable, as – irrespective of the effect of density fluctuations – volume and mass flow should be linked to each other.

Regarding the specific energy demand, a quadratic influence of the shaft rotation speed was observed, which is close to symmetric, relating to the average speed of 80%, reaching its minimum at that point. A possible partial explanation is the attribution of the idle energy demand (as explained above), considering the mass flow curve's shape. Also, faster rotation leads to an increase in the idle energy consumption caused, for example, by tribological effects.

3.3.3. Cutting tool geometry

Considering the mean effects, as shown in Fig. 3, the F-type and XXF-type cutting tools show very similar values, for all investigated responses, except for the mean volume flow, indicating very similar behavior, with differences in the density of the produced material. This interpretation seems feasible, as the XXF cutting-tool was designed to produce a finer product than the F-type tool and is emphasized by the particle size data, shown in Fig. S3 in the supplementary material. Nonetheless, Fig. 4 shows additional differences between the two regarding throughput steadiness for minimum and maximum gap widths. While it is hardly affected by the F-type cutting tool's gap width, it decreases as gap width increases or decreases from the average. As a result, a (not necessarily significantly) more stable throughput was observed for the F-type compared to the XXF-type at the extremes of the gap width.

On the other hand, the V-type tool shows a significantly different behavior compared to the other two. So a change from the F or XXF tool to the V tool has the most substantial observed effect of all factors, independent of the investigated response (except for the shaft rotation speed's influence on the mean mass flow).

While the strength of the effect is affected by the gap width, according to Fig. 4, for the volume flow (increasing strength with increasing gap width) and the throughput steadiness (minimum strength at 50% gap width), the trends are consistent. The V-type shredder has a significantly lower throughput in terms of volume flow (68% of the intercept-value at 50% gap width) and mass flow (83% of the intercept-value). The two values indicate an increase in material density, which is consistent with the optical evaluation of the product, the data in Fig. S3, and the potential effect of the comb system and may partially explain the significantly higher energy demand of the V-shredder (in addition to the lower mass flow).

Moreover, the V-type teeth are higher than those of the other geometries, and the minimum axial gap between the counter comb teeth is smaller. Consequently, higher forces and resulting torques are needed to break particles, according to Feyerer (2020), which increases the required comminution energy and may contribute to lower throughputs due to slowing down of the shaft to raise these higher forces. The findings on the specific energy demand confirm those of Savage et al. (1981), who reported increasing energy demands with decreasing product sizes.

Contrary to the non-favorable observations concerning throughput and energy demand, the V-type has a significantly

higher throughput steadiness than the F and XXF cutting tool. The differences are of a similar order of magnitude, as those reported by Feil and Pretz (2018) between single-shaft and double-shaft shredders. Consequently, the transfer of their observations on different types of shredders than they used needs experimental validation.

The V-type's significant differences can be explained by its unique properties compared to the other two. These are particularly the different shape of its teeth, the larger cutting circle, and the comb system. Most of all, the comb system is expected to contribute to the desirable higher throughput steadiness and the finer material quality, but at the cost of decreasing the mean throughput.

3.3.4. Summary

The derived models provide valuable insights into the influences of shredders' parametrization, which can be directly transferred to industry-scale plants. Based on the cited literature, also assumptions on the physical explanation of the observed effects were discussed. For a more detailed understanding of the effects and for designing new cutting tools, the development of numerical physical models that incorporate solid wastes' specific characteristics is a worthwhile target for future research.

The observations show a high similarity between the F-type and XXF-type cutting tools concerning throughput behavior and energy demand. The observations for the V-type tool, on the other hand, were quite different, with stronger effects than the radial gap width and the shaft rotation speed for almost all responses. The V-type has a more stable but lower throughput, produces finer material, and has a higher specific energy demand.

A larger gap width is favorable in terms of mean throughput and specific energy demand. Regarding throughput steadiness, a medium gap width seems to be best for the XXF tool and worst for the V tool, with hardly any effect on the F tool.

For shaft rotation speed, a turning point was observed regarding mean mass flow, showing the best value at 84%, close to the best value for the specific energy demand at about 80%. Regarding throughput steadiness, lower gap widths are better.

4. Conclusions

Significant models were found concerning the investigated parameters' influence on all regarded responses, with empirical significances < 0.0001. The high p_{lof} (>0.3) additionally confirm the suitability of the used models for describing the trends of the curves, except for the model for the volume flow (p_{lof}-value of 0.0507).

Consequently, it is shown that DoE-based experimentation and analysis is suitable for drawing reliable conclusions from industry-scale mechanical processing experiments with real MSW, despite the wastes' variability. The example case of commercial waste shredding, which was used to demonstrate the method, proves that this is even true for very sophisticated cases, where the waste from one run cannot be re-used in another.

Concerning the shredding of commercial waste, it was shown that the choice of the cutting tool geometry plays an important role when designing mechanical treatment plants for mixed commercial waste. Its influence can hardly be compensated by changing the gap width and shaft rotation speed.

In practice, many additional influences must be considered to find the optimal parametrization of the shredder. For example, the economic optimum depends on processing fees, product and energy prices, the legally approved plant capacity, the capacity of subsequent machines, and the quantitative impact of throughput fluctuations on the overall process's performance.

Ultimately, what was only marginally discussed in this study is the effect of the factors on the produced material's quality in terms of grain size. However, it is highly relevant for the performance of downstream machinery, product quality, and the technical and economic evaluation of, particularly, the high differences regarding mean throughput and energy demand of the V cutting tool, compared to the others. The data for this evaluation were collected during the described experiment, and first insights were provided in the appendix. But a detailed analysis requires extending the method that was successfully demonstrated in this work to responses that are compositional vectors and is subject to future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wasman.2020.12.015.

References

Aarne Vesilind, P., Rimer, A.E., Worrell, W.A., 1980. Performance characteristics of a hammermill shredder. In: Proceedings 1980 National Waste Processing Conference. ASME, Washington, D.C.

Chambers, J.M., Hastie, T.J., 1993. Statistical models. In: Chambers, J.M., Hastie, T.J. (Eds.), Statistical models in S, pp. 13–44.

Cleary, P.W., Morrison, R.D., 2021. Geometric analysis of cone crusher liner shape: Geometric measures, methods for their calculation and linkage to crusher behaviour. Miner. Eng. 160, 106701. https://doi.org/10.1016/j.mineng.2020.106701.

Coskun, E., Bosling, M., Dombrowa, M., Feil, A., Pretz, T., 2017. Mechanische Aufbereitungsprozesse effizient gestalten [Efficient design of mechanical treatment processes]. In: Thomé-Kozmiensky, K.J., Goldmann, D. (Eds.), Recycling und Rohstoffe, vol. 10. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 407–426.

Curtis, A., Küppers, B., Möllnitz, S., Khodier, K., Sarc, R., 2020. Real time material flow monitoring in mechanical waste processing and the relevance of fluctuations. Waste Manage. https://doi.org/10.1016/j.wasman.2020.10.037.

Dean, A., Voss, D., Draguljić, D., 2017. Design and Analysis of Experiments. Springer, Cham, p. 840.

Dong, K., Esfandiary, A.H., Yu, A.B., 2017. Discrete particle simulation of particle flow and separation on a vibrating screen: effect of aperture shape. Powder Technol. 314, 195–202. https://doi.org/10.1016/j.powtec.2016.11.004.

European Union, 2018. Directive (EU) 2018/851 of the European Parliament and of the council of 30 May 2018 amending Directive 2008/98/EC on waste.

Feil, A., Coskun, E., Bosling, M., Kaufeld, S., Pretz, T., 2019. Improvement of the recycling of plastics in lightweight packaging treatment plants by a process control concept. Waste Manage. Res.: J. Int. Solid Wastes Public Cleans. Assoc., ISWA 37 (2), 120–126. https://doi.org/10.1177/0734242X19826372.

Feil, A., Pretz, T., 2018. Ungenutzte Potenziale in der Abfallaufbereitung [Unexploited potentials in waste processing]. In: Pomberger, R. (Ed.), Recy & DepoTech 2018: Vorträge-Konferenzband zur 14. Recy & DepoTech-Konferenz. AVAW Eigenverlag.

Feyerer, C., 2020. Interaktion des Belastungskollektives und der Werkzeuggeometrie eines langsamlaufenden Einwellenzerkleinerers [Interaction of the load collective and the tool geometry of a low-speed single-shaft shredder]. Master's thesis at Montanuniversitaet Leoben.

Greenacre, M., 2019. Compositional Data Analysis in Practice. CRC Press, Boca Raton, Florida, XIV.

Hand, D.J., Taylor, C.C., 1987. Multivariate Analysis of Variance and Repeated Measures. Springer, Netherlands, Dordrecht.

Johnson, R.A., Wichern, D.W., 2007. Applied Multivariate Statistical Analysis. Pearson/Prentice Hall, Upper Saddle River, NJ.

Kaufeld, S., Coskun, E., Bosling, M., Pretz, T., 2017. Siebwirkungsgrad von Trommelsieben in Abfallbehandlungsanlagen: Untersuchung zum Einfluss zweidimensionaler Materialien [Screening effieciency of drum screens in waste processing plants: investigation on the influence of two-dimensional materials]. In: Thomé-Kozmiensky, K.J., Goldmann, D. (Eds.), Recycling und Rohstoffe. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 439–452.

Kazemi, K., Zhang, B., Lye, L.M., Cai, Q., Cao, T., 2016. Design of experiment (DOE) based screening of factors affecting municipal solid waste (MSW) composting. Waste Manage. (New York N.Y.) 58, 107–117. https://doi.org/10.1016/j.wasman.2016.08.029.

Khodier, K., Curtis, A., Sarc, R., Lehner, M., O'Leary, P., Pomberger, R., 2019a. Smart solid waste processing plant: vision and pathway. In: International Solid Waste Association (Ed.), Proceedings of ISWA word congress 2019, 7–9 October, Bilbao, Spain, pp. 194–195.

Khodier, K., Viczek, S.A., Curtis, A., Aldrian, A., O'Leary, P., Lehner, M., Sarc, R., 2020. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis. Int. J. Environ. Sci. Technol. 74 (1), 1. https://doi.org/10.1007/s13762-019-02526-w.

Küppers, B., Schlögl, S., Friedrich, K., Lederle, L., Pichler, C., Freil, J., Pomberger, R., Vollprecht, D., 2020. Influence of material alterations and machine impairment on throughput related sensor-based sorting performance. Waste Manage. Res. 1–8. https://doi.org/10.1177/0734242X20936745.

Lee, H., Kwon, J.H., Kim, K.H., Cho, H.C., 2008. Application of DEM model to breakage and liberation behaviour of recycled aggregates from impact-breakage of concrete waste. Miner. Eng. 21 (11), 761–765. https://doi.org/10.1016/j.mineng.2008.06.007.

Luo, S., Xiao, B., Xiao, L., 2010. A novel shredder for municipal solid waste (MSW): influence of feed moisture on breakage performance. Bioresour. Technol. 101 (15), 6256–6258. https://doi.org/10.1016/j.biortech.2010.02.067.

Möllnitz, S., Khodier, K., Pomberger, R., Sarc, R., 2020. Grain size dependent distribution of different plastic types in coarse shredded mixed commercial and municipal waste. Waste Manage. 103, 388–398. https://doi.org/10.1016/j.wasman.2019.12.037.

Müller, W., Bockreis, A., 2015. Mechanical-biological waste treatment and utilization of solid recovered fuels - state of the art. In: Thomé-Kozmiensky, K.J., Thiel, S. (Eds.), Waste Management, vol. 5. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 321–338.

Pieper, C., Maier, G., Pfaff, F., Kruggel-Emden, H., Wirtz, S., Gruna, R., Noack, B., Scherer, V., Längle, T., Beyerer, J., Hanebeck, U.D., 2016. Numerical modeling of an automated optical belt sorter using the Discrete Element Method. Powder Technol. 301, 805–814. https://doi.org/10.1016/j.powtec.2016.07.018.

Pomberger, R., 2008. Entwicklung von Ersatzbrennstoff für das HOTDISC-Verfahren und Analyse der abfallwirtschaftlichen Relevanz [Development of refuse-derived fuel for the HOTDISC process and analysis of the relevance for waste management]. Dissertation at Montanuniversitaet Leoben.

Qin, Y., Wu, J., Zhou, Q., Xu, Z., 2009. Quadratic nonlinear models for optimizing electrostatic separation of crushed waste printed circuit boards using response surface methodology. J. Hazard. Mater. 167 (1–3), 1038–1043. https://doi.org/10.1016/j.jhazmat.2009.01.088.

Sarc, R., Curtis, A., Kandlbauer, L., Khodier, K., Lorber, K.E., Pomberger, R., 2019a. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management - A review. Waste Manage. 95, 476–492. https://doi.org/10.1016/j.wasman.2019.06.035.

Sarc, R., Lorber, K.E., Pomberger, R., 2016. Manufacturing of solid recovered fuels (SRF) for energy recovery processes. In: Thomé-Kozmiensky, K.J., Thiel, S. (Eds.), Waste Management. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 401–416.

Sarc, R., Seidler, I.M., Kandlbauer, L., Lorber, K.E., Pomberger, R., 2019b. Design, quality and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry - Update 2019. Waste Manage. Res. 37 (9), 885–897. https://doi.org/10.1177/0734242X19862600.

Savage, G.M., Diaz, L.F., Trezek, G.J., Golueke, C.G., Wiles, C., Oberacker, D., 1981. On-site evaluation of municipal solid waste shredders. Resour. Recovery Conserv. 5 (4), 343–362. https://doi.org/10.1016/0304-3967(81)90016-0.

Siebertz, K., van Bebber, D., Hochkirchen, T., 2010. Statistische Versuchsplanung [Design of Experiments]. Springer Berlin Heidelberg, Berlin, Heidelberg.

Sinnott, M.D., Cleary, P.W., 2015. Simulation of particle flows and breakage in crushers using DEM: Part 2 – Impact crushers. Miner. Eng. 74, 163–177. https://doi.org/10.1016/j.mineng.2014.11.017.

Stat-Ease, Inc., 2018. Optimality Criteria. https://www.statease.com/docs/v11/contents/advanced-topics/optimality-criteria/ (accessed: May 26th, 2020).

Umweltbundesamt, 2020. Abfallaufkommen [Waste generation]. https://www.umweltbundesamt.de/daten/ressourcen-abfall/abfallaufkommen (last updated: August 18th, 2020).

Wasserman, L., 2013. All of Statistics: A Concise Course in Statistical Inference. Springer, New York, NY, p. 446.

Weissenbach, T., Sarc, R., 2020. Investigation of particle-specific characteristics of non-hazardous, fine shredded mixed waste. Waste Manage. (New York N.Y.) 119, 162–171. https://doi.org/10.1016/j.wasman.2020.09.033.

Wuschke, L., Jäckel, H.-G., Leißner, T., Peuker, U.A., 2019. Crushing of large Li-ion battery cells. Waste Manage. (New York N.Y.) 85, 317–326. https://doi.org/10.1016/j.wasman.2018.12.042.

Publication III

Distribution-independent empirical modeling of particle size distributions – coarse-shredding of mixed commercial waste

K. Khodier, R. Sarc

Processes, 2021, 9 (3), https://doi.org/10.3390/pr9030414

Author contributions according to the CRediT system:

KK: conceptualization, methodology, software, formal analysis, investigation, data curation, writing – original draft, writing – review & editing, visualization, project administration

RS: writing – review and editing, supervision, project administration, funding acquisition

Article

Distribution-Independent Empirical Modeling of Particle Size Distributions—Coarse-Shredding of Mixed Commercial Waste

Karim Khodier [1] and Renato Sarc [2,*]

1. Department of Environmental and Energy Process Engineering, Chair of Process Technology and Industrial Environmental Protection, Montanuniversitaet Leoben, Franz-Josef-Straße 18, 8700 Leoben, Austria; karim.khodier@unileoben.ac.at
2. Department of Environmental and Energy Process Engineering, Chair of Waste Processing Technology and Waste Management, Montanuniversitaet Leoben, Franz-Josef-Straße 18, 8700 Leoben, Austria
* Correspondence: renato.sarc@unileoben.ac.at

Abstract: Particle size distributions (PSDs) belong to the most critical properties of particulate materials. They influence process behavior and product qualities. Standard methods for describing them are either too detailed for straightforward interpretation (i.e., lists of individual particles), hide too much information (summary values), or are distribution-dependent, limiting their applicability to distributions produced by a small number of processes. In this work the distribution-independent approach of modeling isometric log-ratio-transformed shares of an arbitrary number of discrete particle size classes is presented. It allows using standard empirical modeling techniques, and the mathematically proper calculation of confidence and prediction regions. The method is demonstrated on coarse-shredding of mixed commercial waste from Styria in Austria, resulting in a significant model for the influence of shredding parameters on produced particle sizes (with classes: >80 mm, 30–80 mm, 0–30 mm). It identifies the cutting tool geometry as significant, with a *p*-value $< 10^{-5}$, while evaluating the gap width and shaft rotation speed as non-significant. In conclusion, the results question typically chosen operation parameters in practice, and the applied method has proven to be valuable addition to the mathematical toolbox of process engineers.

1. Introduction

The size distribution belongs to the most critical properties of solid particulate materials, and particularly mixed solid waste, for example: The quality classes of solid recovered fuels (SRF) demand specific maximum particle sizes [1]. The particle size distribution (PSD) of the organic fraction of municipal waste impacts its anaerobic digestion [2]. The particle sizes of municipal solid waste influence the yields of dry gas, char, and tar in fixed bed reactor pyrolysis [3]. And the PSD influences the mass throughput of robotic sorters, which are limited by picks per hour [4]; hence, smaller particle sizes (and the corresponding smaller weights) decrease the possible mass throughput.

Concerning mixed commercial waste, besides the PSD's relevance as a technical quality criterion for processing products (e.g., SRF [5]), and its influence on the performance of reactors and processing machines (e.g., wind sifters [6]), different types of materials also concentrate in different particle size ranges (e.g., sorting analysis by Khodier et al. [7] and the size distribution of different plastic types according to Möllnitz et al. [8]). Hence, beyond influencing the shares of a plant's throughput that pass specific machines (due to material flow separation by screens), the PSD also determines the kinds of materials that pass through these machines.

Therefore, beneficial PSDs increase the effectiveness, as well as economic and ecologic efficiency of mixed solid waste treatment. Consequently, the PSD of mixed solid waste is

deliberately influenced during mechanical processing, which is usually the first treatment stage for this kind of material, mainly through a combination of shredding and sieving [1–3,7,9–13].

1.1. Describing Particle Size Distributions

Strictly speaking, the PSD of a collective is described as a list of the individual particles' sizes. But the representation as such a list is not suitable for analyzing and comparing PSDs. Consequently, various more useful methods for describing PSDs exist, which were summarized by Polke et al. [14]: Collectives of particles are often described through average equivalent diameters. An example is the Sauter diameter d_S, which gives information on the specific surface of the collective (Equation (1), where V_t is the total volume of all particles, and A_t is the total surface area of all particles).

$$d_S = 6V_t / A_t \tag{1}$$

Often, information on the width of the distribution is also essential. Consequently, measures of this width are frequently provided. Examples are the sample standard deviation S (Equation (2), where d_i is the size of the ith particle, $\bar{d}$ is the arithmetic average size, and N is the number of particles), or distribution-independent measures of the width, as shown, for example, in Equation (3) (where W is the width, and d_b is the bth percentile particle size).

$$S = \sqrt{\frac{\sum_{i=1}^{N}\left(d_i - \bar{d}\right)^2}{N - 1}} \tag{2}$$

$$W = \frac{d_{75} - d_{25}}{d_{50}} \tag{3}$$

Considering the described influence of the PSD on the path individual particles take through a plant, it is essential in mechanical waste processing to have more detailed knowledge on it than just summary values. Hence, the results of a PSD analysis are often reported graphically: the frequency density is shown in a histogram [15], where particle sizes are summarized into particle size classes (PSCs)—which is also the level of information obtained from sieve analyses (Figure 1a). Another representation, which is more suitable for comparing PSDs, is the sum distribution (Figure 1b) [14].
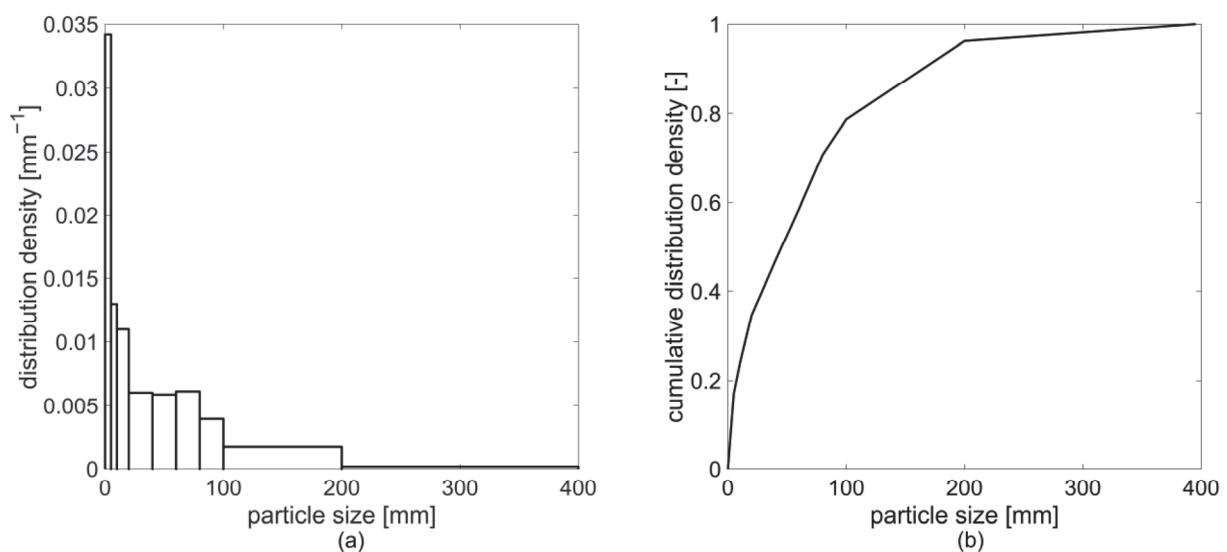

Figure 1. Representation of the overall particle size distribution (PSD) of a mixed commercial waste according to Reference [7]: (**a**) frequency density; (**b**) cumulative frequency density.

These graphical representations correspond to an empirical distribution [15]. As the sample size approaches infinity and the class width approaches zero, the histogram's representation becomes a continuous function: the probability density function (PDF). This PDF can also be approximated from analyses, where only summary information on discrete PSCs is available (e.g., sieve analyses), for example, through cubic splines [16] or Kernel density estimation [17].

Sometimes, the PDF can be approximately described by an analytical expression. In such cases, reporting the momentums of such an analytical distribution is sufficient to describe the PSD. So, for example, reporting the arithmetic mean particle size $\bar{d}$ of the sample and its standard deviation S, usually implies the underlying assumption of a normal distribution, according to Equation (4), where $q(d)$ is the probability density or frequency density for particles of size d, μ is the arithmetic average of the population, which is estimated through $\bar{d}$, and σ is the population's standard deviation, which is estimated through S [18].

$$q(d) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right) \tag{4}$$

Three further analytical PDFs, are reported as being relevant to the description of PSDs [14]: The log-normal distribution describes materials, where the logarithm of the particle size follows a normal distribution. It is a positively skewed distribution, which—in contrast to the normal distribution—only includes positive and, therefore, meaningful particle sizes. It is shown in Equation (5), where μ and σ are estimated by the arithmetic average and the sample standard deviation of the logarithm of the particle sizes.

$$q(d) = \begin{cases} 0 & \text{if } d < 0 \\ \frac{1}{d\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(\log(d)-\mu)^2}{2\sigma^2}\right) & \text{otherwise} \end{cases} \tag{5}$$

The Gates-Gaudin-Schuhmann (GGS) distribution [19] is an empirical approximation that is often suitable for describing products of coarse comminution processes. Its parameters are the maximum particle size d_{max} and the uniformity parameter m_u. Its PDF is shown in Equation (6).

$$q(d) = \begin{cases} 0 & \text{if } d < 0 \text{ or } d > d_{max} \\ \frac{m_u}{d_{max}}\left(\frac{d}{d_{max}}\right)^{m_u-1} & \text{otherwise} \end{cases} \tag{6}$$

The Rosin-Rammler-Sperling-Bennet (RRSB) distribution [20] is also an empirical approximation. Its first parameter d' is equal to the particle size, where the cumulative frequency density reaches the value $1 - 1/e \approx 0.632$, and its second parameter n_u is a uniformity parameter. Its PDF is shown in Equation (7). The RRSB distribution is often used for describing products of fine comminution and dusts. Examples of the discussed distributions are shown in Figure 2.

$$q(d) = \begin{cases} 0 & \text{if } d < 0 \\ \frac{n_u}{d'}\left(\frac{d}{d'}\right)^{n_u-1} \exp\left[-\left(\frac{d}{d'}\right)^{n_u}\right] & \text{otherwise} \end{cases} \tag{7}$$
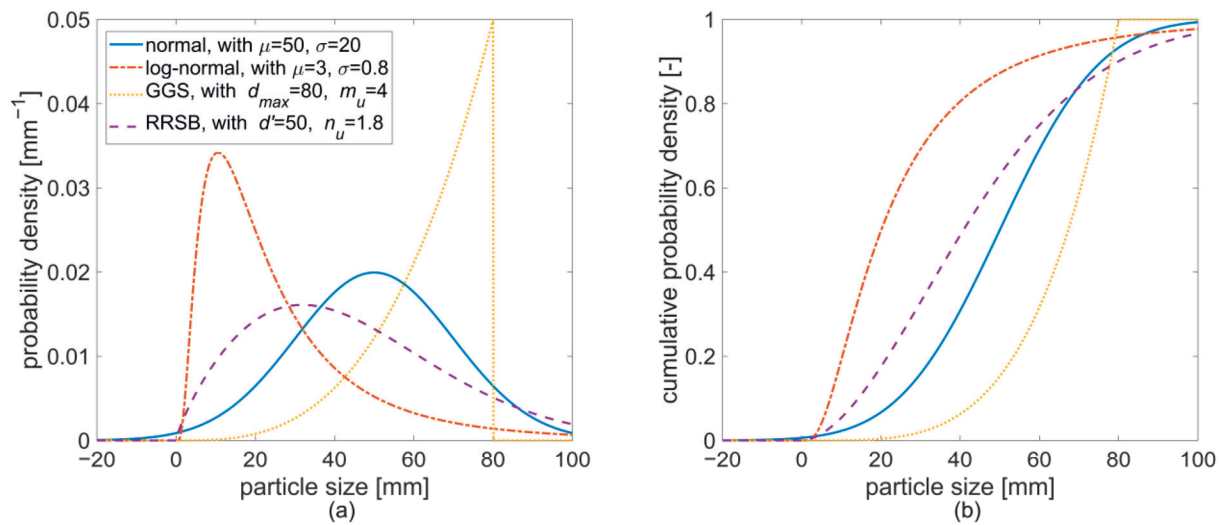
Figure 2. Frequency density (**a**) and cumulative frequency density (**b**) of a normal, log-normal, Gates-Gaudin-Schuhmann (GGS), and Rosin-Rammler-Sperling-Bennet (RRSB) distribution.

1.2. Modeling Particle Size Distributions

Products' PSDs result from their original condition and the kinds and parameters of machines that process them. Hence, to beneficially influence PSDs through process design and the choice and parametrization of machines, modeling and predicting them is desirable.

The most sophisticated and advantageous models are physical models, which provide an in-depth understanding of the phenomena that influence PSDs. Simulations based on such models are usually implemented using the discrete element method (DEM). Examples in literature are the works of Sinnott and Cleary [21] on the particle flows and breakage in impact crushers, Lee et al. [22] on breakage and liberation behavior of recycled aggregates from impact-breakage of concrete waste, and Dong et al. [23] on particle flow and separation on vibrating screens.

While DEM-based models improve process understanding, they also require high amounts of computational resources and detailed models and data on the processing machines and the materials to be comminuted, which limits their applicability in practice in many cases [24]: no published models, for example, incorporate the variability of materials, geometries and particle interactions for real mixed solid waste.

When physical models cannot be used, empirical regression models can deliver basic insights on machine and parameter influences on PSDs. For ensuring the reliability of the results, it is essential to involve statistical analyses for finding and interpreting the models. For scalar target values, the procedure has been thoroughly described by Khodier et al. [24], based on the example of the parametrization-dependent energy demand and throughput behavior of coarse shredders for mixed commercial waste.

PSDs are non-scalar: they are either described as PDFs, which are continuous functions, or as compositions of PSCs and, therefore, as multivariate vectors. Analytical PDFs, as those presented in Section 1.1., are defined by a set of momentums, which can also be treated as multivariate vectors. Consequently, modeling PSDs requires extensions of the methods used by Khodier et al. [24] to multivariate dependent variables.

The most widely applied variation of regression modeling is linear regression. It relates one or more dependent variables to one or more independent variables based on a set of linear regression coefficients. Its most general form—which is relevant to this work—is multivariate multiple linear regression, which involves multiple independent variables and multivariate dependent variables. Its model Equation is shown in Equation (8) [25]. The matrix Y (Equation (9)) is the $P \times R$ matrix of P observations of the R-variate vector

of the dependent variable, with elements $y_{p,r}$. $\boldsymbol{\varepsilon}$ (Equation (12)) is also a $P \times R$ matrix and shows the corresponding model residuals $\varepsilon_{p,r}$. $\boldsymbol{X}$ (Equation (10)) is a $P \times (Q+1)$ matrix of the settings $x_{p,q}$ of the Q independent variables corresponding to the P observations. Its first column (which is indexed with 0) is a column of ones, corresponding to constant terms in linear regressions. And the matrix $\boldsymbol{\beta}$ (Equation (11)) contains the $(Q+1) \times R$ regression coefficients $\beta_{q,r}$.

$$Y = X \cdot \beta + \varepsilon \tag{8}$$

$$Y = \begin{bmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,R} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ y_{P,1} & y_{P,2} & \cdots & y_{P,R} \end{bmatrix} \tag{9}$$

$$X = \begin{bmatrix} x_{1,0} & x_{1,1} & x_{1,2} & \cdots & x_{1,Q} \\ x_{2,0} & x_{2,1} & x_{2,2} & \cdots & x_{2,Q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{P,0} & x_{P,1} & x_{P,2} & \cdots & x_{P,Q} \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,Q} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,Q} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{P,1} & x_{P,2} & \cdots & x_{P,Q} \end{bmatrix} \tag{10}$$

$$\beta = \begin{bmatrix} \beta_{0,1} & \beta_{0,2} & \cdots & \beta_{0,R} \\ \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,R} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{Q,1} & \beta_{Q,2} & \cdots & \beta_{Q,R} \end{bmatrix} \tag{11}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_{1,1} & \varepsilon_{1,2} & \cdots & \varepsilon_{1,R} \\ \varepsilon_{2,1} & \varepsilon_{2,2} & \cdots & \varepsilon_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{P,1} & \varepsilon_{P,2} & \cdots & \varepsilon_{P,R} \end{bmatrix} \tag{12}$$

The resulting model is obtained by minimizing the sum of squares of the residuals $\varepsilon_{p,r}$. It is shown in Equation (13), where $\hat{Y}$ is the matrix of the $P \times R$ model predictions $\hat{y}_{p,r}$ for the dependent variable, corresponding to the observations in Y, and $\hat{\beta}$ is the matrix of the $(Q+1) \times R$ least-squares estimates $\hat{\beta}_{q,r}$ of the regression coefficients in β.

$$\hat{Y} = X \cdot \hat{\beta} \tag{13}$$

For linear regression models, it is necessary to describe the dependent variable as a vector of a fixed length. In the case of analytical PDFs, this is the vector of the momentums. So, the immediate result of a model for these momentums is the vector of their expected values, and the corresponding confidence bands. From there, the PDF and its confidence region can be calculated.

Theoretically, $\hat{\beta}$ can also be calculated from Q univariate linear regressions. The resulting values $\hat{\beta}_{q,r}$ are identical. But multivariate methods, involving multivariate linear regression and multivariate analysis of variance (MANOVA, e.g., Reference [26]), are preferable: they allow a more accurate calculation of confidence regions, considering correlations between the momentums. Moreover, the evaluation of parameter's significance should be based on multivariate considerations to find coherent models for the resulting—interdependent—distribution of the particle sizes.

Analytical PDFs are practical, as they allow a detailed description of PSDs with a small number of variables—the momentums. However, only few processes produce PSDs that follow analytical functions, according to Polke et al. [14]. Consequently, methods for distribution-independent modeling of PSDs are also needed, which is especially true for the mechanical processing of mixed solid waste, considering the variety of processing machines used there: e.g., shredders, screens, magnetic separators, and sensor-based sorters [10].

Empirical distributions allow the distribution-independent description of PSDs: particle sizes are amalgamated to D discrete PSCs. As a result, the PSD is described as a composition—a D-dimensional vector of the shares of those PSCs.

Mathematically, the compositional nature of the vector of PSCs has wide-reaching consequences: D-dimensional compositions are constrained to a vector space called the D-dimensional simplex $S^{(D)}$, which is a sub-space of the D-dimensional real space $R^{(D)}$ [27]. Compositions, being simplicial vectors, have specific common properties: Their D parts are not linearly independent. This results from a summation constraint: their parts sum up to a fixed constant, e.g., 1 or 100%. Furthermore, parts may only have positive values or zero. So, more precisely, $S^{(D)}$ is a sub-space of $R_0^{+(D)}$. The most well-known representation of a simplex is the ternary diagram (see, e.g., Reference [28]), which shows the three-dimensional simplex.

When modeling simplicial vectors, the constraints of the simplex and the interdependence of the compositional parts must be considered. While Khodier et al. [29] report the empirical observation that the constraints are automatically fulfilled for the prediction values $\hat{Y}$, if all observations in Y are valid compositions, this does not apply to corresponding confidence regions. Hence, different methods were developed in the past decades to handle and model simplicial data.

The most widely applied approach for handling compositions is transforming data using log-ratios, as suggested by Aitchison [30]. There are many kinds of such log-ratios, but the state of the art approach in the compositional data community is the application of so-called isometric log-ratios (ilr) (as proposed by Egozcue [31]), according to Reference [32]. These are bijective projections of the D-dimensional Simplex onto a $(D-1)$-dimensional real-space with an orthonormal basis [27]. The resulting ilr-coordinates are unconstrained, linearly independent coordinates. Hence, standard statistics can be applied in the projected log-ratio space, as demonstrated, for example, by Edjabou et al. [33] for waste composition analysis. Consequently, the use of ilr-transformations allows the application of standard multivariate multiple linear regression to predict PSDs, described as empirical distributions.

In this work, the approach of modeling influences on empirical PSDs using multivariate multiple linear regression and ilrs is applied on particle size data, using the programming language R. The data was obtained within Khodier et al.'s [24] industry-scale coarse-shredding experiments with real mixed commercial waste. Based on this example, this work aims at presenting the method to the process engineering and waste processing communities, enabling the distribution-independent empirical modeling of particle size distributions, while preserving all relevant information. Moreover, the method's suitability for PSD modeling and its limitations and potential pitfalls in the interpretation of the results are discussed. And, finally, insights on the influences of coarse shredders' parameters on the PSDs of mixed commercial waste are presented, complementing findings of Khodier et al. [24] on their effects on shredders' throughput behavior and energy demand.

2. Materials and Methods

2.1. Experimental Design and Setup

The choice of the experimental design and the shredding experiment setup have been explained in detail by Khodier et al. [24]. Hence, they are only summarized in the following. The extension of the experiment with material sampling and particle size analysis are described in detail.

2.1.1. Experimental Design

The shredding experiment examines the influence of three independent variables on the throughput behavior and energy demand of a Terminator 5000 SD, which is a single-shaft shredder from the Austrian company Komptech GmbH (Frohnleiten, Austria)—a research partner in the funded project ReWaste 4.0. These independent variables are the radial gap width w, the shaft rotation speed s, and the cutting tool geometry c. The factor range of w was defined from 0% to 100% of the maximum gap width, with discrete levels

at a step size of 10%. The minimum of the factor range of *s* was chosen at 60% and the maximum at 100% of the maximum shaft rotation speed of 31 rpm, again width discrete 10% steps. Concerning *c*, three different geometries are examined, called "F", "XXF", and "V" (for more details, cf. Reference [24] and Figure A1 and Table A1 in the Appendix A).

The factors are coded for the design and analysis of the experiment: for the numerical factors *s* and *w*, their range is adjusted to −1 to 1, representing the minimum and maximum factor settings. Concerning the nominal factor *c*, it is represented by two variables c_1 and c_2, based on sum contrasts (cf. Reference [34]). The values of these variables that correspond to the cutting tool geometries are shown in Table 1.

Table 1. Contrast matrix for the cutting tool geometry.

Cutting Tool Geometry	c_1	c_2
F	1	0
XXF	−1	−1
V	0	1

The experimental design settings of the independent variables were chosen based on a statistical Design of Experiments (cf. Reference [35]). A 32 runs, completely randomized D-optimal design (cf. Reference [36]) was chosen, with no blocking (cf. Reference [37]) and with five replicate points and five lack-of-fit points. It is based on the reduced-cubic design model, shown in Equation (14), where $\hat{y}^{(r)}$ is the model prediction for the rth (univariate) response (=dependent variable), $\vec{x}$ is a vector of the factors' settings, and $K_{jkmn}^{(r)}$ is the model coefficient for the rth response and the factor or interaction (=multiplication of factors) $w^j s^k c_1^m c_2^n$.

$$\hat{y}^{(r)}\left(\vec{x}\right) = \sum_{j=0}^{2}\sum_{k=0}^{2-j}\sum_{m=0}^{1}\sum_{n=0}^{1-m}\left(K_{jkmn}^{(r)} w^j s^k c_1^m c_2^n\right) \tag{14}$$

Equation (14) can easily be extended to multivariate responses; therefore, the design is also valid for multivariate multiple linear regression. To extend it, $\hat{y}^{(r)}$ equals $\hat{y}_{p,r}$ in Equation (13): the rth univariate response becomes the rth dimension of the multivariate response, and *p* indexes the corresponding vector of factor settings $\vec{x}$, which is simply a row of *X*. Each factor or interaction $w^j s^k c_1^m c_2^n$ is represented by a column of *X*, and each coefficient $K_{jkmn}^{(r)}$ corresponds to a coefficient $\hat{\beta}_{q,r}$.

2.1.2. Setup of the Shredding Experiment

The flowchart of the experiment is shown in Figure A2 in the Appendix A: The feed material is waste, declared as mixed commercial waste (cf. Reference [24] for more details), collected in Styria in Austria in October 2019. It is fed into the shredder's feeding bunker using a wheel loader. From the shredder's output belt, the material is passed to a digital material flow monitoring system (DMFMS), consisting of a belt-scale and optical sensors (cf. Reference [11]). The material leaving the DMFMS is collected on a product heap. Each experimental run has a total duration of one hour.

2.1.3. Sampling

For analyzing the PSD of the shredded waste, samples must be taken. The design of the sampling process in this experiment is based on Pierre Gy's Theory of Sampling (TOS), as described in the Danish standard DS 3077 [38] and the work of Khodier et al. [7] on its application on coarsely shredded mixed commercial waste.

According to TOS, the fundamental sampling principle must be considered: each particle must have the same probability of ending up in the sample. The most beneficial sampling situation is a one-dimensional sampling [39], e.g., taking the samples from a falling stream. Furthermore, a sample should spatially cover the whole lot, which is

achieved by composite sampling: the sample is a composite of so-called increments (the sampled material from one individual sampling step).

In this work, the sample for each experimental run consisted of 40 such increments. They were taken from a falling stream, swiveling the output conveyor belt of the DMFMS back and forth over a sampling box, elevated by a forklift (see Figure 3), every 3 minutes, taking two increments. The inner dimensions of the box were 115 × 915 × 565 (length × width × height in mm). For ensuring that all desired material ends up in the box, the height of the back-side wall of the box was increased by placing 1.5 m-long wood boards inside. The box was changed after every ten increments. Based on the share of the samples in the total processed mass, each increment covered approximately 0.61 seconds of throughput. For intermediate storage, the samples were finally transferred to 1 m^3 big bags.

Figure 3. Sampling setup.

2.1.4. Particle Size Analysis

The PSDs of the samples were analyzed using a Komptech (Frohnleiten, AT) Nemus 2700 drum screen and five screening drums that have square holes with side lengths of 80, 60, 40, 20, and 10 mm. The geometries of the drums are shown in Figure A3 and Table A2 in the Appendix A.

First, one big bag of a sample was evenly distributed in the feeding bunker of the screen, which has a length of 4033 mm in the direction of the material flow, and a width of 1035 mm. Then, the drum was started with a rotation speed of 11.5 rpm, and the conveyor belt of the feeding bunker was operated at 0.026 m/s. The drum screen was only stopped after all material had passed. The produced fine fraction was then screened with the subsequent finer drum. The scale used for measuring the masses of the fraction has an uncertainty of 10 g.

2.2. *Analysis of the Results*

The analyses of the results, which are explained in this section, were performed in R version 4.0.2, based on the work of van den Boogaart and Tolosana-Delgado [40]. The implementation is attached as an HTML export of a jupyter notebook (see Supplementary Material).

For the analyses in this work, the six particle size fractions from the particle size analysis are aggregated to three PSCs for easier visualization. To ensure the relevance of the findings for mechanical waste processing, the PSCs were chosen, based on the particle

size limits of SRF premium quality (30 mm) and SRF medium quality (80 mm) [1]. Since none of the used screen drums had a mesh width of 30 mm, equal shares of the screening fraction 20–40 mm are assigned to the particle size fractions 0–30 mm and 30–80 mm, which corresponds to linear interpolation.

2.2.1. Isometric Log-Ratios

The ilr transformation (denoted as a function ilr()) is an isometry of the vector spaces S^D and R^{D-1} [27], which means that the distance between two compositions $\vec{y}^{(1)}$ and $\vec{y}^{(2)}$ is preserved in the transformation. For interpretation, it is essential to understand that the distance referred to is not the Euclidian distance Δ_E (Equation (15), where y_i is the ith element of a R-dimensional composition $\vec{y}$), but rather the Aitchison distance Δ_A (Equation (16), where $ilr_i()$ is the function for the ith ilr dimension) (cf. Reference [27]). Consequently, the preserved distance is of a relative, multiplicative nature and not an absolute, additive nature. Hence, the least-squares minimization of the model residuals in ε, when calculating the coefficients' estimates $\hat{\beta}$, is also based on the Aitchison distance when applying the ilr transformation. Considering this is particularly important when the order of magnitude of the parts' shares differs significantly since small absolute differences become very significant on a relative scale for very small shares (cf. Reference [41]).

$$\Delta_E\left(\vec{y}^{(1)}, \vec{y}^{(2)}\right) = \sqrt{\sum_{i=1}^{R}\left(y_i^{(1)} - y_i^{(2)}\right)^2} \tag{15}$$

$$\Delta_A\left(\vec{y}^{(1)}, \vec{y}^{(2)}\right) = \sqrt{\frac{1}{2R}\sum_{i=1}^{R}\sum_{j=1}^{R}\left[\ln\left(\frac{y_i^{(1)}}{y_j^{(1)}}\right) - \ln\left(\frac{y_i^{(2)}}{y_j^{(2)}}\right)\right]^2} = \sqrt{\sum_{i=1}^{R-1}\left[ilr_i\left(\vec{y}^{(1)}\right) - ilr_i\left(\vec{y}^{(2)}\right)\right]^2} \tag{16}$$

Furthermore, the ilr transformation is not defined if any part has a value of zero. While there are different approaches to handling such values, they complicate the application of the transformation [27].

In the following, ilr-transformed coordinates are marked with "$*$" so that $Y^* = ilr(Y)$. The back-transformation function is denoted "$ilr^{-1}()$". For calculating ilr coordinates, the compositional parts are sequentially grouped: first, each component is assigned to a group +1, or −1. For subsequent ilr coordinates, the elements of one group are again assigned to new groups +1 or −1, and the parts of the other group are assigned to group 0. Each ilr coordinate is then calculated according to Equation (17), where $a_{i,r}$ is the scaling factor for the rth compositional part and the ith ilr coordinate. The calculation of $a_{i,r}$ is shown in Equation (18), where t is the number of parts in group +1, and v is the number of parts in group −1 [31].

$$ilr_i\left(\vec{y}\right) = \sum_{r=1}^{R}\left[a_{i,r}\ln(y_r)\right] \tag{17}$$

$$a_{i,r} = \begin{cases} +\frac{1}{t}\sqrt{\frac{tv}{t+v}} & \text{for parts of group } +1 \\ -\frac{1}{v}\sqrt{\frac{tv}{t+v}} & \text{for parts of group } -1 \\ 0 & \text{for parts of group } 0 \end{cases} \tag{18}$$

Greenacre [42] documents concerns regarding the interpretation of ilr-transformed data since it incorporates the geometric means of compositional parts. Consequently, in this work, results are interpreted based on graphical representations of back-transformed data. Hence, the exact choice of groups is arbitrary, and the standard grouping of the "ilr()" function of the "compositions" package version 2.0-1 in R [43] is used. In this work, $r = 1$ stands for the fraction > 80 mm, $r = 2$ for the fraction 30–80 mm, and $r = 3$ for the fraction 0–30 mm (see Supplementary Material). Resulting from the standard grouping,

the ilr-transformed representation $\vec{y}^{*}$ of a particle size composition $\vec{y}$ (corresponding to a row of Y) is calculated according to Equation (19).

$$\vec{y}^{*} = \mathrm{ilr}\left(\vec{y}\right) = \begin{pmatrix} \ln\left[\left(\dfrac{y_2}{y_1}\right)^{\sqrt{0.5}}\right] \\ \ln\left[\dfrac{y_3^{\sqrt{2/3}}}{(y_1 y_2)^{\sqrt{1/6}}}\right] \end{pmatrix} \tag{19}$$

2.2.2. Model Reduction: MANOVA

To obtain a final, reliable model, the factors and interactions in Equation (14) must be checked on their significance, eliminating non-significant ones, but retaining model hierarchy (cf. Reference [24]). For univariate dependent variables, this is done through F-tests in an analysis of variance (ANOVA) (cf. Reference [35]). The multivariate character of the compositional dependent variable in this work requires a multivariate extension of the ANOVA: the MANOVA. Different from the ANOVA, there is more than one definition of the F-statistic in the MANOVA. The most commonly used definitions are the Pillai-Barlett trace, Wilk's lambda, the Hotelling-Lawley trace, and Roy's largest eigenvalue statistic, according to Hand and Taylor [26]. These are also the ones implemented in the "regr" package, version 1.1 [44], used for the MANOVA in this work.

For the analyses at hand, the Pillai-Barlett trace was chosen, based on the recommendation of Olson [45] as cited in Reference [26]. Model reduction is performed applying backward selection: the least significant term, which can be removed without violating model hierarchy, is eliminated, as long as removable non-significant factors or interactions are present (cf. Reference [40]). Analogous to Reference [24], 0.1 is chosen as the threshold so that factors and interactions with an empirical significance (*p*-value) higher than that threshold are discarded. The relevant *p*-values are calculated using the "drop1()" function from the "regr" package (version 1.1).

For evaluating the final model, three performance values are calculated: The coefficient of determination R^2 calculates how much of the variance of the data is explained by the model. The adjusted coefficient of determination R^2_{adj} is a measure similar to R^2. But it is adjusted by the terms in the model and thereby evaluates the model's efficiency [35]. Both are calculated using the "R2()" function in R.

The prediction coefficient of determination R^2_{pred} determines the share of variance, which is explained by models fitted without considering the very point which is evaluated. High differences between R^2_{pred} and R^2_{adj} indicate overfitting. R^2_{pred} is calculated according to Equation (20), where $PRESS$ is the prediction residual sum of squares and SS_{tot} is the total sum of squares [46]. $PRESS$ is calculated, using the "PRESS()" function from the "MPV" package version 1.56 [47]. And SS_{tot} is calculated, according to Equation (21). $y^{*}_{p,r}$ is the pth observation of the rth of $(R-1)$ coordinates of the ilr-transformed dependent variable. And $\bar{y}^{*}_r$ is the arithmetic mean of the observations of the rth ilr coordinate (see Equation (22)).

$$R^2_{pred} = 1 - \frac{PRESS}{SS_{tot}} \tag{20}$$

$$SS_{tot} = \sum_{p=1}^{P} \sum_{r=1}^{R-1} \left(y^{*}_{p,r} - \bar{y}^{*}_r\right)^2 \tag{21}$$

$$\bar{y}^{*}_r = \frac{1}{P} \sum_{p=1}^{P} y^{*}_{p,r} \tag{22}$$

2.2.3. Analysis of the Residuals

The tests in the MANOVA require multivariate normality of the (ilr-transformed) residuals ε^{*}. Hence, to validate the final model, the distribution of the residuals must be

examined. Each coordinate of variables, which follow a multivariate normal distribution, also follows a univariate normal distribution [48]. Hence, a quantile-quantile plot for each coordinate is examined as a first visual step. Since the individual coordinates' univariate normality is a necessary, but not sufficient condition for multivariate normality, multivariate tests are also applied, particularly Mardia's Skewness and Mardia's Kurtosis [49]. Both tests are part of the "mvn()" function in "MVN" package version 5.8 [50], which also tests the univariate normality of the individual coordinates using the Shapiro-Wilk test.

2.2.4. Confidence and Prediction

The resulting model Equation shows the most likely prediction value. Two regions express the uncertainty of this prediction: the confidence region and the prediction region (cf. Reference [40]). The confidence region covers the uncertainty of the model parameter estimation. The region reflects likely average PSC distributions (on an ilr scale) for specific parameter settings, for extended operation times, on a chosen confidence level. The prediction region adds the residual variability around the expected value. Hence, it shows likely PSC distributions for one hour of operation (since this is the experimental duration the data is based on).

Due to the required multivariate character of the ilr-transformed residuals, the resulting confidence and prediction regions are equipotential-ellipses (resulting from the PDF of the multivariate normal distribution) on an ilr-scale. Van den Boogaart and Tolosana Delgado [40] provide R-code for calculating these regions and visualizing their back-transformed representation in ternary diagrams. This code is used in this work (see Supplementary Materials).

3. Results and Discussion
3.1. Data and Model

The experimental design and the resulting shares of the PSCs are shown in the Supplementary Materials. Their order corresponds to the order in which the experimental runs were performed. Originally, a completely randomized order was planned. As reported by Khodier et al. [24], due to an unintentional change of the motor rotation speed of the V cutting tool during the experiments, three runs had to be repeated. Since the tight timescale did not allow re-randomizing all remaining runs, considering the time consumed when switching shredders, the randomness is slightly impaired.

A significant model was found based on the experimental data. It is visualized in Figure 4 and shown in Equation (23), where $\hat{\vec{y}}$ is a vector showing the model predictions of the mass shares of the three PSCs. Its first element corresponds to the coarse fraction (>80 mm), its second element to the medium fraction (30–80 mm), and the third element shows the fine fraction (0–30 mm). As shown in the Equation, the only significant factor (at a threshold of $p = 0.1$) is the cutting tool geometry c.

$$\hat{\vec{y}} = \mathrm{ilr}^{-1}\left[\begin{pmatrix} 0.137 \\ 0.047 \end{pmatrix} - \begin{pmatrix} 0.307 \\ 0.178 \end{pmatrix}c_1 + \begin{pmatrix} 0.344 \\ 0.145 \end{pmatrix}c_2\right] = \begin{cases} \begin{pmatrix} 0.39 & 0.31 & 0.30 \end{pmatrix}^T & \text{for } c = \text{"F"} \\ \begin{pmatrix} 0.30 & 0.35 & 0.35 \end{pmatrix}^T & \text{for } c = \text{"XXF"} \\ \begin{pmatrix} 0.21 & 0.42 & 0.37 \end{pmatrix}^T & \text{for } c = \text{"V"} \end{cases} \quad (23)$$

The analysis of the ilr-residuals proves a multivariate normal distribution, with p-values of 0.67 and 0.89 for the Shapiro-Wilk test on univariate normality of each coordinate and p-values of 0.89 and 0.71 for Mardia's Skewness and Mardia's Kurtosis, respectively.

R^2 for the model is 0.57 and R^2_{adj} is 0.54. These values are noticeably lower than those of the models for the throughput behavior and energy demand (see Reference [24]), which range from 0.73 to 0.87 for R^2 and from 0.67 to 0.81 for R^2_{adj}.

The lower coefficients of determination are reasonable since material fluctuations are likely to influence the material quality more than the process behavior, and sampling adds random noise (cf. Reference [7]). Considering the high expectable noise in processing

experiments with real waste, and especially in shredding experiments (cf. Reference [24]), the model performance is satisfactory.
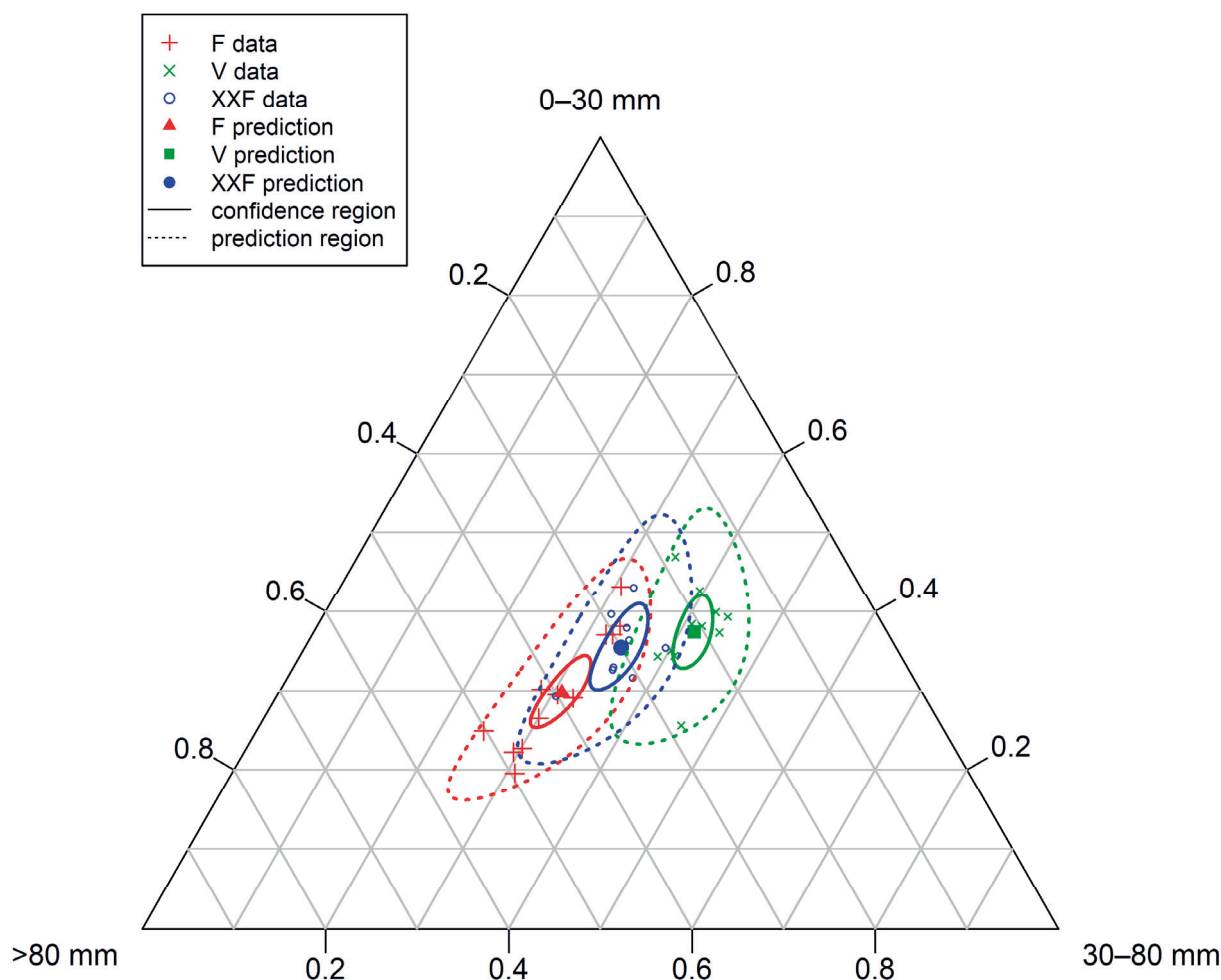
Figure 4. Prediction values and confidence and prediction regions for the particle size class distributions resulting from different cutting tool geometries.

3.2. Discussion of the Method

The data from the particle size analyses were amalgamated to three discrete PSCs and then freed from the restrictions of the simplex by applying an ilr-transformation. The resulting model, the multivariate normal distribution of its residuals, and the consequent successful calculation of confidence and prediction regions prove the potential of this approach in terms of distribution-independent, mathematically correct empirical modeling of particle size distributions.

It is another potential application of ilr-transformations in the context of waste management, besides the application on waste compositions, in terms of sorting fractions (cf. Reference [33]). The method is a notable extension to the toolbox of chemical and process engineers in general, besides familiar approaches, like equivalent diameters or certain analytical distributions (see Section 1.1).

The discussed restriction of the method concerning zero-values is not an issue for the present data. When data include zeros, the potential impact of zero-handling techniques must be considered. These include amalgamations of compositional parts or zero-replacement (cf. Reference [42]).

Furthermore, the impact of the relative scale of ilr-transformed data must be kept in mind. Concerning the aim of the present work, it cannot be ignored, since the absolute

values of the PCSs' shares result in absolute waste masses, which are of interest. But the effect of the relative scale is low when the shares of all parts are of similar orders of magnitude, as is the case in this work (see Equation (23)). For other cases, calculations of weighted variances (cf. Reference [42]), for example, can be used to counteract the impact of the relative scale, if necessary. The relative scale can also be beneficial in other cases (as explained by Pawlowsky-Glahn et al. [27]), for example, where the share of trace elements in a chemical composition has a high impact.

The discussed issue of interpreting ilrs, is solved by graphical representation in this work. While this is a straightforward approach for three or fewer fractions, the question of how to represent more fractions arises. Graphic solutions include sets of ternary diagrams of two specific fractions and amalgamations of the others (cf. Reference [40]) and area plots (for the expected values, but not for confidence and prediction regions). Some other approaches include: The application of easier interpretable log-ratios (e.g., additive log-ratios or log-ratios incorporating amalgamations, cf. Reference [42]) at the cost of the exact representation of the variance structure in the data, or modeling non-transformed data, while incorporating potential issues with the restrictions of the simplex, particularly for calculating confidence and prediction regions.

Finally, the results in Figure 4 confirm the decision of Khodier et al. [24] to perform a Design of Experiments-based investigation, with multiple runs: the 95% prediction regions overlap even for the F and V unit. Hence, conclusions that contradict the findings in this work could be drawn when comparing only single runs of one hour.

3.3. Discussion of the Modeling Results

For the gap width and shaft rotation speed, no significant impacts on the produced PSDs were identified. In conclusion, either no such effects exist, or they are too small, compared to the residual variance in the data, to be detected based on the data at hand (resulting in so-called type II or β error). According to Biemann [51], no matter how small, any effect becomes significant if the amount of data is big enough. Hence, the order of magnitude of potential, non-identified effects is of interest. For the chosen limit *p*-value of 0.1 in the MANOVA, linear effects are preserved in the model if the 90% confidence regions of their extreme settings (which get smaller with more data) do not overlap. And potential effects are likely not to exceed the maximum distance of the borders of these regions.

Consequently, the confidence regions for the minimum and maximum settings of *w* and *s* at factor setting 0 for the other variables were plotted (see Figure 5), based on a linear model (see Equation (24)), to give a visual impression of potential type II errors. The residuals of these models also follow a multivariate normal distribution. Figure 5 shows, that *w* influences the allocation to PSCs of 0 to about 13% of the product material, and *s* influences 0 to about 9%, for average settings of the other factors.

$$
\hat{\vec{y}} = \mathrm{ilr}^{-1} \left[\begin{pmatrix} K^{(1)}_{0000} \\ K^{(2)}_{0000} \end{pmatrix} + \begin{pmatrix} K^{(1)}_{1000} \\ K^{(2)}_{1000} \end{pmatrix} w + \begin{pmatrix} K^{(1)}_{0100} \\ K^{(2)}_{0100} \end{pmatrix} s + \begin{pmatrix} K^{(1)}_{0010} \\ K^{(2)}_{0010} \end{pmatrix} c_1 + \begin{pmatrix} K^{(1)}_{0001} \\ K^{(2)}_{0001} \end{pmatrix} c_2 \right] \tag{24}
$$

Differences in the PSD, caused by the shaft rotation speed, would most likely be based on differences in the breakage of brittle materials in the waste, caused by the impact speed of the shaft's teeth. Since the used machines are slow running shredders, the non-existence of such effects appears feasible. Furthermore, since potential impacts of this kind are most likely relatively small, the low share of brittle materials in mixed commercial waste (cf. Reference [7]) complicates their identification.

Concerning the gap width, considering the geometry of the cutting tools, a slight increase of the coarse fraction (>80 mm) would be reasonable due to falling through of uncomminuted particles between the teeth of the counter comb. But the non-significance of potential effects makes sense, considering that the breakage situation, according to Feyerer [52], does not change with the gap width for the F and XXF geometry and only slightly for the V geometry.

The influence of the cutting tool geometry is highly significant, with a factor *p*-value $< 10^{-5}$. As Figure 4 shows, it is largest, comparing the PSDs produced by the F and the V geometry. But the F and XXF geometries also differ significantly on a 95% confidence level. The results confirm expectations, considering the geometries (cf. Figure A1 and Table A1 in the Appendix A): The smaller axial gap between the counter comb teeth of the XXF geometry leads to a finer product, compared to the F unit. To be more precise, the share of the coarse fraction decreases in favor of the two other fractions.

Concerning the V geometry, the gap between the counter comb teeth is smaller than the gap of the F geometry. It is larger than the XXF geometry's gap close to the shaft but gets much smaller with increasing distance. This smaller gap, combined with a comb system (cf. Reference [24]), leads to the finest product among the examined geometries—again, mainly in terms of an even lower share of coarse material compared to the XXF geometry.

Relating the results to the findings of Khodier et al. [24], the choice of a cutting tool geometry depends on the requirements of the process: The V tool produces the finest material of the three but at the cost of higher energy demand and smaller but steadier throughput.

Concerning the gap width and the shaft rotation speed, the standard operation with minimum gap widths and maximum shaft rotation speeds must be questioned: increasing the gap width is beneficial for the throughput and energy demand and hardly affects the throughput steadiness, according to Reference [24]. The shares of the PSCs chosen in this work (typical for SRF production) are not or only a little affected, with a maximum-likelihood influence on only about 3% of the material.

For the shaft rotation speed, the maximum likelihood influence of this non-significant factor on the PSD only concerns about 2% of the material, while the mass flow and energy demand show an optimum at about 84% and 80% of the maximum shaft rotation speed, respectively.
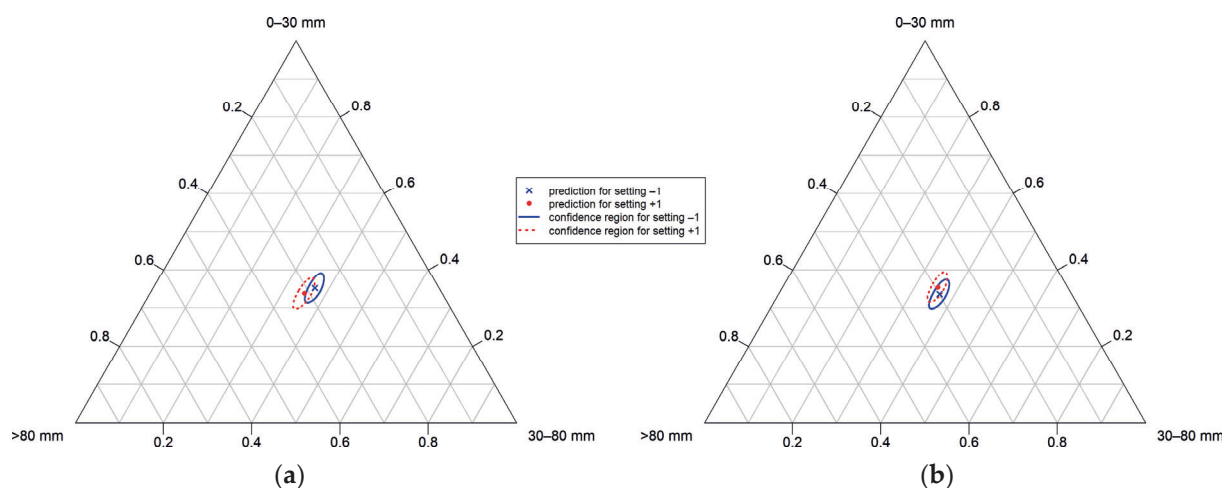
Figure 5. Ninety percent confidence regions for the minimum (−1) and maximum (+1) settings of *w* (**a**) and *s* (**b**), at factor setting 0 for the corresponding other factors of *w*, *s*, c_1 and c_2, based on the linear model from Equation (24).

4. Conclusions

Multivariate multiple linear regression modeling was applied in this work on ilr-transformed PSC data from a Design of Experiments-based 32 runs coarse-shredding experiment with mixed commercial waste. A significant model, with an R^2 of 0.57 was found, identifying the cutting tool geometry as a highly significant influence on the PSD.

The gap width and shaft rotation speed were found not to be significant, with maximum-likelihood influences on 3% and 2% on the material, respectively. If the discussed potential type II errors are rated as economically relevant or other particle size classes are of interest, further data should be generated and analyzed. Otherwise, the PSD can be

treated as invariant to these factors when optimizing the throughput and energy demand. Consequently, based on the new insights from this work, a much more efficient operation of mechanical waste processing plants can be reached. The influence of the cutting tool, on the contrary, is highly significant. Its choice depends on process requirements.

What was not investigated are selective influences on specific material fractions, e.g., metals or wood. The investigation of the PSDs of such fractions may give more detailed insights and lead to different conclusions on process parametrization. It is subject to further research, which can make use of the presented modeling methods.

Using the presented method requires an in-depth understanding of the implications of applying the ilr transformation. It is essential to avoid wrong interpretations, caused by the introduced relative scale or by zero-replacement practices, where necessary. Nonetheless, establishing such understanding is rewarding: in conclusion, the method has proven to be suitable for the distribution-independent modeling of PSDs. Hence, it is a valuable addition to the toolkit of engineers, dealing with particulate materials.

Abbreviations

$a_{i,r}$	scaling factor for the rth compositional part and the ith ilr coordinate
ANOVA	analysis of variance
A_t	total surface area of all particles
c	cutting tool geometry
c_1, c_2	coded representation of the cutting tool geometry
d	particle size
$\overline{d}$	arithmetic average particle size
d'	characteristic particle size of the RRSB distribution
d_b	b^{th} percentile particle size
d_i	size of the ith particle
d_{max}	maximum particle size
d_S	Sauter diameter
D	number of particle size classes
DEM	discrete element method
DMFMS	digital material flow monitoring system
GGS	Gates-Gaudin-Schuhmann
ilr	isometric log-ratios
ilr_i	i^{th} ilr dimension
j	factor exponent
k	factor exponent

$K_{jkmn}^{(r)}$	model constant for the factor or interaction $w^j s^k c_1^m c_2^n$ and the response r
m	factor exponent
m_u	uniformity parameter of the GGS distribution
MANOVA	multivariate analysis of variance
n	factor exponent
n_u	uniformity parameter of the RRSB distribution
N	number of particles
p	empirical significance
P	number of observations
PDF	probability density function
$PRESS$	prediction residual sum of squares
PSC	particle size class
PSD	particle size distribution
$q(d)$	frequency density for particles of size d
Q	number of independent variables
R	number of dimensions of the dependent variable
R^2	coefficient of determination
R_{adj}^2	adjusted coefficient of determination
R_{pred}^2	prediction coefficient of determination
$\boldsymbol{R}^{(D)}$	D-dimensional real space
$\boldsymbol{R}_0^{+(D)}$	D-dimensional positive real space, including 0
RRSB	Rosin-Rammler-Sperling-Bennet
s	shaft rotation speed
S	sample standard deviation
$\boldsymbol{S}^{(D)}$	D-dimensional simplex
SRF	solid recovered fuel
SS_{tot}	total sum of squares
t	number of parts in group $+1$
TOS	Theory of Sampling
v	number of parts in group -1
V_t	total volume of all particles
w	radial gap width
W	width of a distribution
$x_{p,q}$	p^{th} setting of the qth independent variable
$\boldsymbol{X}$	matrix of settings of the independent variables
$\vec{y}$	compositional vector
$\vec{y}^*$	ilr-transformed compositional vector
$\hat{\vec{y}}$	model prediction of the shares of the particle size classes
y_i	i^{th} element of $\vec{y}$
$y_{p,r}$	p^{th} observation of the rth dimension of the dependent variable
$y_{p,r}^*$	p^{th} observation of the rth dimension of the ilr-transformed dependent variable
$\hat{y}_{p,r}$	model prediction for $y_{p,r}$
$\hat{y}^{(r)}$	model prediction of the response r
$\overline{y}_r^*$	arithmetic mean of the rth ilr coordinate
$\boldsymbol{Y}$	matrix of the dependent variable
$\boldsymbol{Y}^*$	matrix of the ilr-transformed dependent variable
$\hat{\boldsymbol{Y}}$	matrix of model predictions of the dependent variable
$\boldsymbol{\beta}$	matrix of the regression coefficients
$\hat{\boldsymbol{\beta}}$	matrix of least squares estimates of the regression coefficients
$\beta_{q,r}$	q^{th} regression coefficient for the rth dimension of the dependent variable
$\hat{\beta}_{q,r}$	least squares estimate of $\beta_{q,r}$
Δ_A	Aitchison distance
Δ_E	Euclidian distance
$\boldsymbol{\varepsilon}$	matrix of the model residuals
$\boldsymbol{\varepsilon}^*$	matrix of the ilr-transformed residuals
$\varepsilon_{p,r}$	model residual corresponding to $y_{p,r}$
μ	arithmetic average of a population
σ	standard deviation of a population

Appendix A
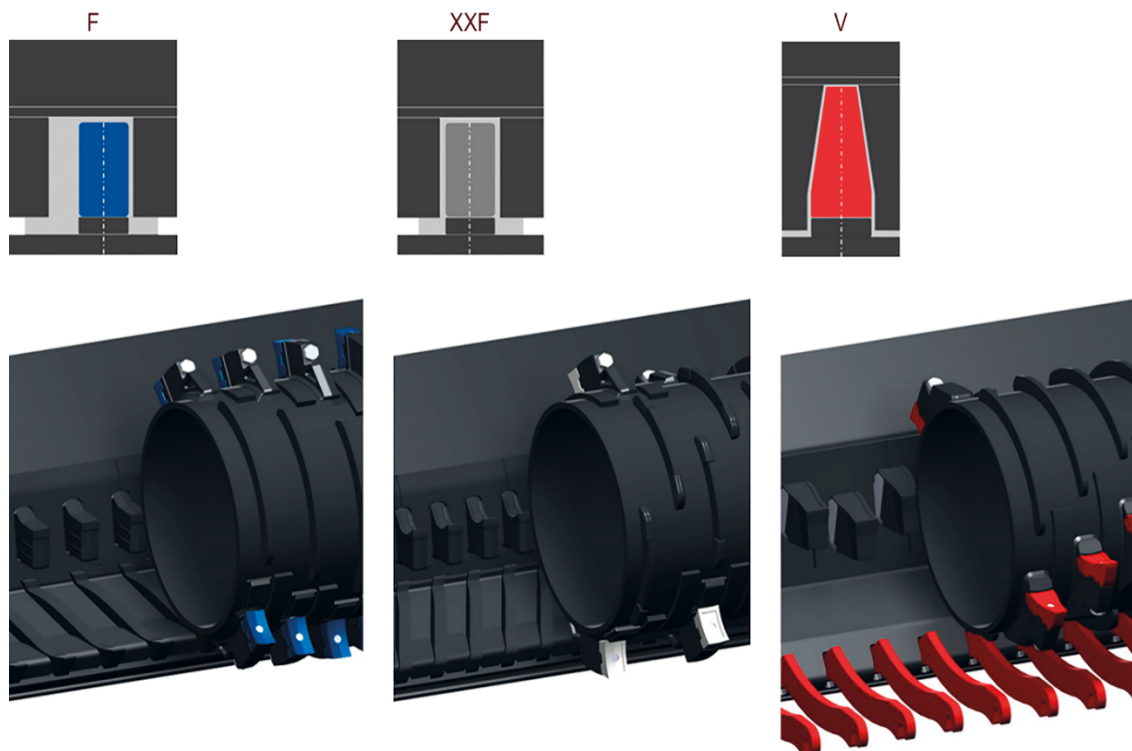
Figure A1. Cutting tool geometries [24].

Table A1. Technical data of the cutting tool geometries [24].

Type	F	XXF	V
number of cutting teeth (shaft) [pcs.]	32	22	32
position of cutting teeth (shaft) [-]	double helix	chevron	chevron
width of cutting teeth (shaft) [mm]	70	70	42/85 *
height of cutting teeth (shaft) [mm]	124	124	183
width of cutting teeth (counter comb) [mm]	64	54	81/100 *
height of cutting teeth (counter comb) [mm]	142	136	202
cutting circle [mm]	1070	1070	1170
length of shredding-shaft [mm]		3000	
right side cutting gap (axial) [mm]	3.5	2	3
left side cutting gap (axial) [mm]	39	2	3
minimum cutting gap (radial) [mm]		0	
maximum cutting gap (radial) [mm]	33	35	30/38 *
comb-system [-]	no	no	yes

* bottom/top of the teeth.

Figure A2. Experimental setup: photo and flow chart [24].
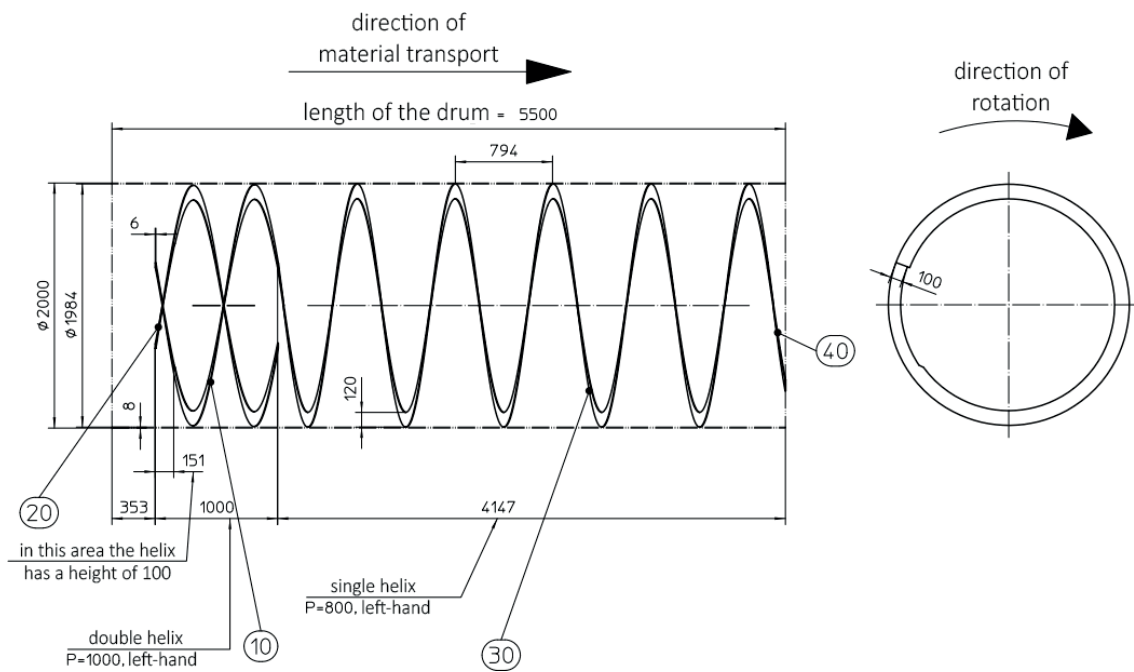
Figure A3. Screening drum geometry.

Table A2. Data of the screening drums.

side length of the square-shaped holes (mm)	80	60	40	20	10
total hole area (m^2)	16, 61	17, 06	17, 14	17, 96	14, 55

References

1. Sarc, R.; Lorber, K.E.; Pomberger, R. Manufacturing of Solid Recovered Fuels (SRF) for Energy Recovery Processes. In *Waste Management*; Thomé-Kozmiensky, K.J., Thiel, S., Eds.; TK Verlag Karl Thomé-Kozmiensky: Neuruppin, Germany, 2016; pp. 401–416.
2. Zhang, Y.; Banks, C.J. Impact of different particle size distributions on anaerobic digestion of the organic fraction of municipal solid waste. *Waste Manag.* **2013**, *33*, 297–307. [CrossRef]

3.　Luo, S.; Xiao, B.; Hu, Z.; Liu, S.; Guan, Y.; Cai, L. Influence of particle size on pyrolysis and gasification performance of municipal solid waste in a fixed bed reactor. *Bioresour. Technol.* **2010**, *101*, 6517–6520. [CrossRef] [PubMed]

4.　Sarc, R.; Curtis, A.; Kandlbauer, L.; Khodier, K.; Lorber, K.E.; Pomberger, R. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management—A review. *Waste Manag.* **2019**, *95*, 476–492. [CrossRef]

5.　Sarc, R.; Seidler, I.M.; Kandlbauer, L.; Lorber, K.E.; Pomberger, R. Design, quality and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry—Update 2019. *Waste Manag. Res.* **2019**, *37*, 885–897. [CrossRef]

6.　Leschonski, K. Windsichten [wind sifting]. In *Handbuch der Mechanischen Verfahrenstechnik [Handbook of Mechanical Process Engineering]*; Schubert, H., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012; pp. 584–612. ISBN 3-527-30577-7.

7.　Khodier, K.; Viczek, S.A.; Curtis, A.; Aldrian, A.; O'Leary, P.; Lehner, M.; Sarc, R. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: Procedure, particle size and sorting analysis. *Int. J. Environ. Sci. Technol.* **2020**, *17*, 959–972. [CrossRef]

8.　Möllnitz, S.; Khodier, K.; Pomberger, R.; Sarc, R. Grain size dependent distribution of different plastic types in coarse shredded mixed commercial and municipal waste. *Waste Manag.* **2020**, *103*, 388–398. [CrossRef]

9.　Feil, A.; Coskun, E.; Bosling, M.; Kaufeld, S.; Pretz, T. Improvement of the recycling of plastics in lightweight packaging treatment plants by a process control concept. *Waste Manag. Res.* **2019**, *37*, 120–126. [CrossRef] [PubMed]

10.　Gundupalli, S.P.; Hait, S.; Thakur, A. A review on automated sorting of source-separated municipal solid waste for recycling. *Waste Manag.* **2017**, *60*, 56–74. [CrossRef] [PubMed]

11.　Curtis, A.; Küppers, B.; Möllnitz, S.; Khodier, K.; Sarc, R. Digital material flow monitoring in waste processing—The relevance of material and throughput fluctuations. *Waste Manag.* **2021**, *120*, 687–697. [CrossRef] [PubMed]

12.　Möllnitz, S.; Küppers, B.; Curtis, A.; Khodier, K.; Sarc, R. Influence of pre-screening on down-stream processing for the production of plastic enriched fractions for recycling from mixed commercial and municipal waste. *Waste Manag.* **2021**, *119*. [CrossRef]

13.　Müller, W.; Bockreis, A. Mechanical-Biological Waste Treatment and Utilization of Solid Recovered Fuels—State of the Art. In *Waste Management*; Thomé-Kozmiensky, K.J., Thiel, S., Eds.; TK Verlag Karl Thomé-Kozmiensky: Neuruppin, Germany, 2015; pp. 321–338.

14.　Polke, R.; Schäfer, M.; Scholz, N. Charakterisierung disperser Systeme [Characterization of disperse systems]. In *Handbuch der Mechanischen Verfahrenstechnik [Handbook of Mechanical Process Engineering]*; Schubert, H., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012; pp. 7–129. ISBN 3-527-30577-7.

15.　Wasserman, L. *All of Statistics: A Concise Course in Statistical Inference*; Springer: New York, NY, USA, 2013; ISBN 978-0-387-21736-9.

16.　Micula, G.; Micula, S. *Handbook of Splines*; Springer Netherlands: Dordrecht, The Netherlands, 1999; ISBN 978-94-010-6244-2.

17.　Scott, D.W. *Multivariate Density Estimation*; Wiley: New York, NY, USA, 1992; ISBN 9780471547709.

18.　Heumann, C.; Michael Schomaker, S. *Introduction to Statistics and Data Analysis: With Exercises, Solutions and Applications in R*; Springer: Berlin/Heidelberg, Germany, 2017; ISBN 978-3319461601.

19.　German Institute for Standardization. *DIN 66143:1974-03, Darstellung von Korn-(Teilchen-)Größenverteilungen; Potenznetz [Graphical Representation of Particle Size Distributions; Power-Function Grid]*; Beuth Verlag GmbH: Berlin, Germany, 1974.

20.　German Institute for Standardization. *DIN 66145:1976-04, Darstellung von Korn-(Teilchen-)Größenverteilungen; RRSB-Netz [Graphical Representation of Particle Size Distributions; RRSB-Grid]*; Beuth Verlag GmbH: Berlin, Germany, 1976.

21.　Sinnott, M.D.; Cleary, P.W. Simulation of particle flows and breakage in crushers using DEM: Part 2—Impact crushers. *Miner. Eng.* **2015**, *74*, 163–177. [CrossRef]

22.　Lee, H.; Kwon, J.H.; Kim, K.H.; Cho, H.C. Application of DEM model to breakage and liberation behaviour of recycled aggregates from impact-breakage of concrete waste. *Miner. Eng.* **2008**, *21*, 761–765. [CrossRef]

23.　Dong, K.; Esfandiary, A.H.; Yu, A.B. Discrete particle simulation of particle flow and separation on a vibrating screen: Effect of aperture shape. *Powder Technol.* **2017**, *314*, 195–202. [CrossRef]

24.　Khodier, K.; Feyerer, C.; Möllnitz, S.; Curtis, A.; Sarc, R. Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: Coarse-shredding of commercial waste. *Waste Manag.* **2021**, *121*, 164–174. [CrossRef]

25.　Johnson, R.A.; Wichern, D.W. *Applied Multivariate Statistical Analysis*, 6th ed.; Pearson/Prentice Hall: Upper Saddle River, NJ, USA, 2007; ISBN 978-0-13-187715-3.

26.　Hand, D.J.; Taylor, C.C. *Multivariate Analysis of Variance and Repeated Measures*; Springer Netherlands: Dordrecht, The Netherlands, 1987; ISBN 978-94-010-7913-6.

27.　Pawlowsky-Glahn, V.; Egozcue, J.J.; Tolosana-Delgado, R. *Modeling and Analysis of Compositional Data*; John Wiley & Sons Inc.: Chichester, UK, 2015; ISBN 9781118443064.

28.　Pomberger, R.; Sarc, R.; Lorber, K.E. Dynamic visualisation of municipal waste management performance in the EU using Ternary Diagram method. *Waste Manag.* **2017**, *61*, 558–571. [CrossRef]

29.　Khodier, K.; Lehner, M.; Sarc, R. Multilinear modeling of particle size distributions. In *Proceedings of the 8th International Workshop on Compositional Data Analysis (CoDaWork2019): Terrassa, 3–8 June 2019*; Ortego, M.I., Ed.; Universitat Politècnica de Catalunya-BarcelonaTECH: Barcelona, Spain, 2019; pp. 82–85. ISBN 978-84-947240-1-5.

30.　Aitchison, J. The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. B* **1982**, *44*, 139–177. [CrossRef]

31.　Egozcue, J.J. Isometric Logratio Tranformations for Compositional Data Analysis. *Math. Geol.* **2003**, *35*, 279–300. [CrossRef]

32.　Weise, D.R.; Palarea-Albaladejo, J.; Johnson, T.J.; Jung, H. Analyzing Wildland Fire Smoke Emissions Data Using Compositional Data Techniques. *J. Geophys. Res. Atmos.* **2020**, *125*, 139. [CrossRef]

33. Edjabou, M.E.; Martín-Fernández, J.A.; Scheutz, C.; Astrup, T.F. Statistical analysis of solid waste composition data: Arithmetic mean, standard deviation and correlation coefficients. *Waste Manag.* **2017**, *69*, 13–23. [CrossRef]

34. Chambers, J.M.; Hastie, T.J. Statistical Models. In *Statistical Models in S*; Chambers, J.M., Hastie, T.J., Eds.; Chapman & Hall: London, UK, 1993; pp. 13–44. ISBN 978-0-534-16765-3.

35. Siebertz, K.; van Bebber, D.; Hochkirchen, T. *Statistische Versuchsplanung [Design of Experiments]*; Springer: Berlin/Heidelberg, Germany, 2010; ISBN 978-3-642-05492-1.

36. Stat-Ease Inc. Optimality Criteria. Available online: https://www.statease.com/docs/v11/contents/advanced-topics/optimality-criteria/ (accessed on 26 May 2020).

37. Dean, A.; Voss, D.; Draguljić, D. *Design and Analysis of Experiments*, 2nd ed.; Springer: Cham, Switzerland, 2017; ISBN 978-3-319-52250-0.

38. Danish Standards Foundation. *DS 3077 Representative Sampling—Horizontal Standard*; Danish Standards Foundation: Charlottenlund, Denmark, 2013.

39. Esbensen, K.H.; Wagner, C. Theory of sampling (TOS) versus measurement uncertainty (MU)—A call for integration. *TrAC Trends Anal. Chem.* **2014**, *57*, 93–106. [CrossRef]

40. van den Boogaart, K.G.; Tolosana-Delgado, R. *Analyzing Compositional Data with R*; Springer: Dordrecht, The Netherlands, 2013; ISBN 978-3-642-36809-7.

41. Khodier, K.; Lehner, M.; Sarc, R. Empirical modeling of compositions in chemical engineering. In *Proceedings of the 16th Minisymposium Verfahrenstechnik and 7th Partikelforum (TU Wien, Sept. 21/22, 2020)*; Jordan, C., Ed.; TU Wien: Vienna, Austria, 2020; pp. 118–121. ISBN 978-3-903337-01-5.

42. Greenacre, M. *Compositional Data Analysis in Practice*; CRC Press: Boca Raton, FL, USA, 2019; ISBN 978-1-138-31643-0.

43. van den Boogart, K.G.; Tolosana-Delgado, R. Package 'Compositions' (Version 2.0-1). Available online: https://cran.r-project.org/web/packages/compositions/compositions.pdf (accessed on 13 January 2021).

44. Stahel, W.A. Package Regr for an Augmented Regression Analysis. Available online: https://rdrr.io/rforge/regr/f/inst/doc/regr-description.pdf (accessed on 14 January 2021).

45. Olson, C.L. On choosing a test statistic in multivariate analysis of variance. *Psychol. Bull.* **1976**, *83*, 579–586. [CrossRef]

46. Pareto, A. Predictive R-Squared According to Tom Hopper. Available online: https://rpubs.com/RatherBit/102428 (accessed on 14 January 2021).

47. Braun, W.J.; MacQueen, S. Package 'MPV' (Version 1.56). Available online: https://cran.r-project.org/web/packages/MPV/MPV.pdf (accessed on 14 January 2021).

48. Wang, C.-C. A MATLAB package for multivariate normality test. *J. Stat. Comput.* **2014**, *85*, 166–188. [CrossRef]

49. Mardia, K.V. Measures of multivariate skewness and kurtosis with applications. *Biometrika* **1970**, *57*, 519–530. [CrossRef]

50. Korkmaz, S.; Goksuluk, D.; Zararsiz, G. Package 'MVN' (Version 5.8). Available online: https://cran.r-project.org/web/packages/MVN/MVN.pdf (accessed on 14 January 2021).

51. Biemann, T. Logik und Kritik des Hypothesentests [Logic and criticism of the hypothesis test]. In *Methodik der Empirischen Forschung [Methodology of Empirical Research]*; Albers, S., Klapper, D., Konradt, U., Walter, A., Wolf, J., Eds.; Springer Fachmedien: Wiesbaden, Germany, 2007; ISBN 978-3-8349-0469-0.

52. Feyerer, C. Interaktion des Belastungskollektives und der Werkzeuggeometrie Eines Langsamlaufenden Einwellenzerkleinerers [Interaction of the Load Collective and Tool Geometry of a Low-Speed Single-Shaft Shredder]. Master's Thesis, Montanuniversitaet Leoben, Leoben, Austria, 2020.

3.2 Online particle size measurement

Publication IV

Sensor-based particle-size determination of shredded mixed commercial waste based on two-dimensional images

L. Kandlbauer, **K. Khodier**, D. Ninevski, R. Sarc

Waste Management, 2021, 120, 784–794, https://doi.org/10.1016/j.wasman.2020.11.003

Author contributions according to the CRediT system:

LK: methodology, software, formal analysis, investigation, data curation, writing – original draft, writing & review and editing, visualization, project administration

KK: conceptualization, methodology, investigation, writing – original draft, writing – review & editing, supervision, project administration

DN: methodology, writing – review & editing

RS: writing – review & editing, supervision, funding acquisition

Sensor-based Particle Size Determination of Shredded Mixed Commercial Waste based on two-dimensional Images

L. Kandlbauer [a], K. Khodier [b], D. Ninevski [c], R. Sarc [a,*]

[a] Department of Environmental and Energy Process Engineering, Chair of Waste Processing Technology and Waste Management, Montanuniversitaet Leoben – Franz-Josef-Straße 18, 8700 Leoben, Austria
[b] Department of Environmental and Energy Process Engineering, Chair of Process Technology and Industrial Environmental Protection, Montanuniversitaet Leoben – Franz-Josef-Straße 18, 8700 Leoben, Austria
[c] Department Product Engineering; Chair of Automation, Montanuniversitaet Leoben – Peter-Tunner-Straße 25/II, 8700 Leoben, Austria

ARTICLE INFO

ABSTRACT

To optimize output streams in mechanical waste treatment plants dynamic particle size control is a promising approach. In addition to relevant actuators – such as an adjustable shredder gap width – this also requires technology for online and real-time measurements of the particle size distribution. The paper at hand presents a model in MATLAB® which extracts information about several geometric descriptors – such as diameters, lengths, areas, shape factors – from 2D images of individual particles taken by RGB cameras of pre-shredded, solid, mixed commercial waste and processes this data in a multivariate regression model using the Partial Least Squares Regression (PLSR) to predict the particle size class of each particle according to a drum screen. The investigated materials in this work are lightweight fraction, plastics, wood, paper-cardboard and residual fraction. The particle sizes are divided into classes defined by the screen cuts (in mm) 80, 60, 40, 20 and 10. The results show assignment reliability for certain materials of over 80%. Furthermore, when considering the results for determining a complete particle size distribution – for an exemplary real waste – the accuracy of the model is as good as 99% for the materials wood, 3D-plastics and residual fraction for each particle size class respectively as assignment errors partially compensate each other.

1. Introduction

The concept of a Smart Waste Factory Network 4.0 (SWFN4.0) is a research vision for mechanical waste treatment plants. According to the definition from (Sarc et al., 2019a), Smart Waste Factory Network means:

"The SWFN4.0 describes a system consisting of several waste treatment plants, which perform different tasks in the waste management system and are interconnected via data streams and logistics systems (e.g. sorting plants, production plants for Solid Recovered Fuels, etc.). The individual processes and machines within the plants as well as the individual plants are digitally connected with each other. This connection of the individual machines and systems and the real-time analysis of the waste streams enable dynamic process control and various actuator systems actively intervene in the processes. In addition, people can cooperate interactively with the technology around them.".

Such Smart Waste Factory concepts and plants will support further waste management developments and the reaching of higher sorting and recycling rates for valuable waste particles available in mixed non-hazardous waste streams. The recycling status in European municipal waste management is given in Pomberger et al. (2017) and as it can be seen, that technical developments are required to reach the high recycling targets set up by the European Commission within the "Circular Economy Package 2018" (European Commission, 2018).

In mechanical waste treatment plants, a shredder is usually the first machine for the treatment of solid municipal and commercial waste (Sarc et al., 2019b; Sarc et al., 2014; Sarc and Lorber, 2013). Together with the properties of the input material it affects the particle sizes of the materials and thus influences the efficiency

---

of all subsequent machines, such as screens, magnetic separators, or sensor-based sorting machines. To beneficially influence the particle sizes in real-time – e. g. to keep them as constant as possible regardless of the variability of the input material – three things are required according to Khodier et al. (2019): optimization algorithms (for example based on artificial neural networks), controllable actuators such as the gap width or the speed of the shaft rotation of the shredding unit, as addressed for example by Khodier et al. (unpublished), as well as real-time measurements of the particle size distribution, where the latter being the focus of this work.

For an output stream characterisation in waste treatment plants, material composition and particle size distribution are basic information. The present situation in real plants is, that this data is determined when the material has already left the plant, which does not allow further manipulation of the quality. Peddireddy et al. (2015) investigated an image analysis tool for characterising the composition of waste-derived fuels in a test set-up. Additionally, methods that use 3D laser triangulation and special techniques of image analysis to detect objects on a conveyor or from bulk (Flamme et al., 2018; Kontny, 2017; Zhang et al., 2013) give information regarding particle measurements, however, neither of the mentioned authors used the methods on mixed (non-hazardous) waste materials. Another approach for the characterisation of material streams is to determine measurements and descriptors of particles – such as projected area/circumference, Feret diameters, shape factors, bounding shapes – from two-dimensional images which is presented in this contribution. This approach was also tested on coal particles by Zhang et al. (2013).

To enable particle size analysis in real-time, machines must be equipped with the necessary sensors/cameras to record image data. It is necessary to create prediction models for particle sizes that use for example image analysis, mathematical or statistical methods to evaluate image data by software. In order to evaluate the effect on the particle size distribution the models must be combined with the information regarding individual particle weights. Krämer (2017) already investigated material-specific surface weights for Solid Recovered Fuels. This approach could be used in combination with the prediction models to produce true screening lines.

To describe the size and shape of irregularly formed objects different options are possible. On one hand, projected area and circumference are common measurements, whereas the indication of meaningful lengths or diameters due to the irregularity of the outlines often turns out to be problematic. For example, in the field of process engineering equivalent diameters are often used (e.g. diameter of a circle of the same area, Sauter diameter, Martin diameter (Yang, 2003)). Additionally, it is possible to describe a two-dimensional object through geometrical shapes (e.g. circle, triangle, rectangle), which enclose the particle with the smallest possible area. A special type of diameter, which is often used for screening, is the Feret diameter, which is defined by the distance between two parallel tangents that fully enclose the particle (principle of a calliper) (Yang, 2003). Also, dimensionless coefficients like roundness or ratios between the projected area of the object and the areas of bounding shapes are possible options to describe particle shapes and sizes (Dunnu et al., 2006; Hentschel and Page, 2003; Olson, 2011; Podczeck, 1997; Zhang et al., 2013; Zimmermann, 1998; Zlatev, 2005). Another approach to describe shapes is by applying Elliptical Fourier Transformation. This method has been in use for decades and uses Fourier descriptors to describe ellipses to approximate the closed contour of a shape based on a chain-code (Kuhl and Giardina, 1982).

This paper presents results which were gained by using the regression method partial least squares (PLS). Here, this method was chosen because multiple, correlated variables are present to predict the corresponding particle size class of each particle. PLS extracts orthogonal factors that further allow building a regression model and identifies latent variables that are linear combinations from the original variables. These latent variables fulfil the criterion of maximal covariance between the explanatory (X) and response (Y) variables and allow modelling Y as a function of X. If Y consists of a single vector of data, the method is known as PLS1 if it represents a table of data with two or more variables PLS2 can be applied as well, but might not lead to better results in all cases (Vandeginste et al., 1998a). PLS enables one to find a small number of factors to fit the data. Here, the choice of the considered number of components affects the ability of prediction of the model and must be chosen individually for each investigation. (Massart et al., 1998) A high number of components will usually benefit the prediction of the data, but increases the risk of overfitting (fitting the model to the noise in the data). Therefore, the ratio of components to data points must be well considered. To examine for overfitting, k-fold cross-validation (Vandeginste et al., 1998a) is a proposed method for choosing the correct number of components, which doesn't use the same data for the calibration of the model and the error prediction. Here, the data set is randomly split into k calibration steps ($5 < k < 15$), where each subset of data is tested on the model that was built from the remaining sets. k is often chosen to be five or ten, which is based on experience (Kuhn and Johnson, 2016).

The paper at hand presents methods to generate image material of individual objects with known particle size- and material class, and software that was developed in MATLAB® 2019 to process and analyse the image data. The software calculates particle descriptors from binary images, which were transformed from RGB images, to predict the particle size class through a regression model. The outcomes of the regression model are combined with a theoretical approach applied to a realistic waste composition to show the accuracy of the model.

2. Materials and methods

Here, the methods to generate binary image data from particles with assigned information regarding material and particle size are described. Additionally, the used ways to calculate particle descriptors, as well as the applied regression models, are presented. A visualisation of the individual steps is shown in Fig. 1.

2.1. Preparation of material

The work subsequently presents results that are obtained from images of individual particles from coarsely shredded solid mixed commercial waste which was collected in Styria (AUT) in October 2019. The processing of the material was performed with industrially sized machinery from the company Komptech and included a shredding step with following particle size classification through a screening process and additional manual sorting into six material classes. In total, 32 individual, representative samples were used to generate fractions with defined particle size and material class.

The shredding process was performed through single-shaft shredding units (Komptech Terminator 5000 SD with cutting units F, V and XXF, described in detail by Khodier et al. (unpublished)) with different settings regarding shaft rotation speed and width of the radial cutting gap. This ensured the influence of different settings to be considered in the results. The screening was carried out with a drum screen (Komptech Nemus 3000), where cylindrical drums with 2 m in diameter and 5.5 m in length were used. All used drums had round holes with the diameters (in mm) 80, 60, 40, 20 and 10. To approximate ideal screening, a maximum volume of 1 m³ of material was used per batch. This amount was spread out on the conveyor belt of the feeding bunker of the screen over a length of 4 m. The speed of the conveyor was set to 0.026 m/s, while the rotation speed of the drum was set to 10 rpm. The sorting process involved the assignment for each particle to one of the following material classes: lightweight fraction (e. g. foils, 2D-plastics), plastics (3D-plastics), wood, metal, paper-cardboard and residual fraction.

To reduce the effort for manual sorting due to the large number of samples, machine-aided separation was used to gain the light-
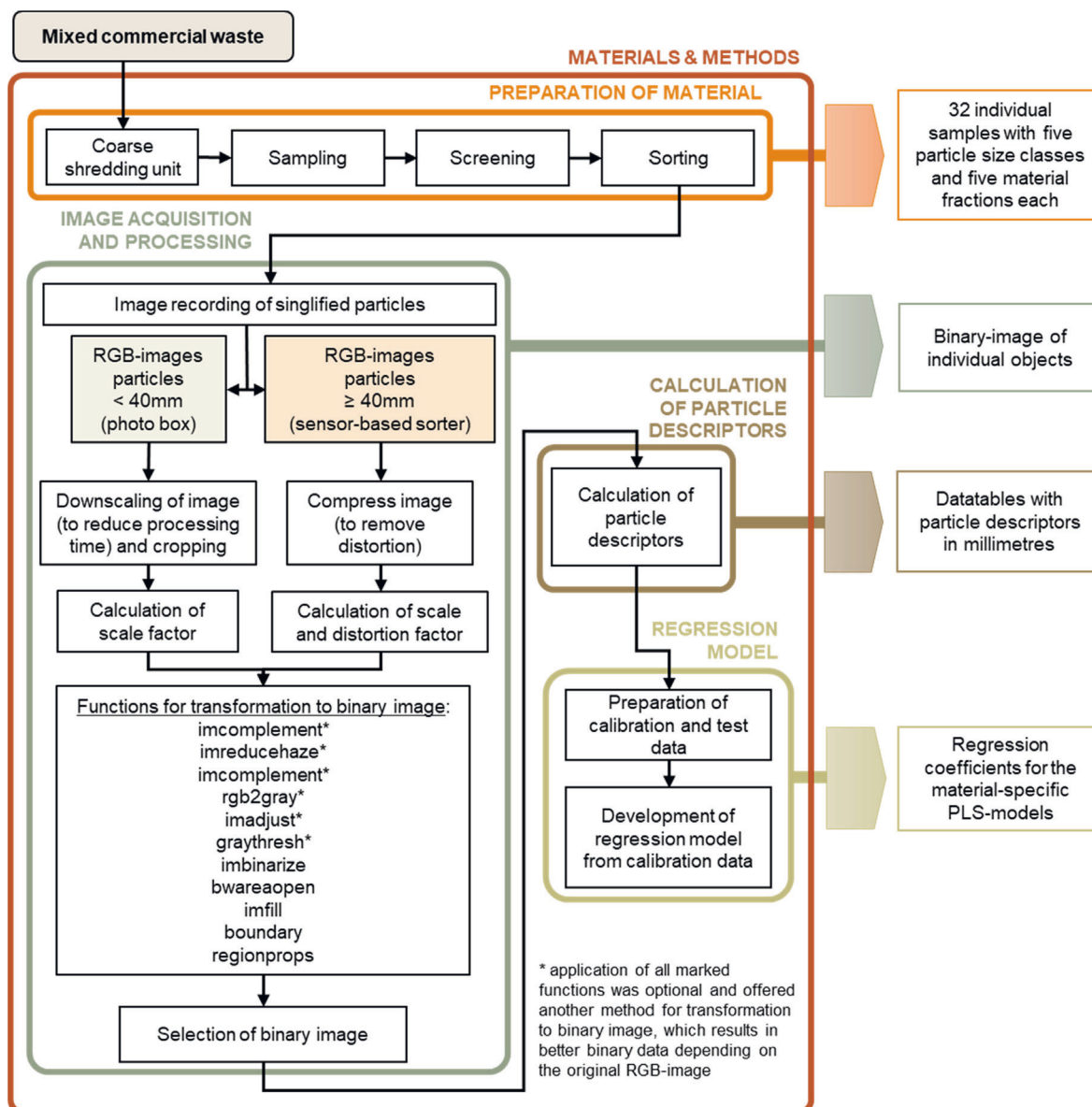

Fig. 1. Schematic layout of the individual applied steps with an overview of the used MATLAB® functions (*) for the image-processing step.

weight fraction. Here, a perforated tube with air nozzles was attached in the expansion cover of a sensor-based sorting machine. The tube was of the length of the working width of the sorting machine and fed via compressed air. The lightweight fraction was separated from the rest of the material stream based on the principle of an air classifier. Due to this, the lightweight fraction contained mainly two-dimensional material like plastic foils, but also light objects of styrofoam and foamed plastic. Manual post-sorting of the material was carried out after the separation to make sure each object was assigned to the correct fraction. Hence, a manual evaluation was still the final criterion for sorting decisions.

2.2. Image acquisition and processing

The developed software requires images of single objects. In this work, this criterion was met by photographing each object individually. Due to technical and time limitations, two different approaches (image acquisition methods) for the recording of the images were applied. Particles assigned to particle sizes $\geq$40 mm were recorded by an RGB-camera in *.png format from the sensor-based sorting machine Redwave 2i. Depending on the particle size, the images were roughly 200x250 pixels in size with a resolution of 2.403 mm^2/pixel. Here, the camera was mounted over the accelerating conveyor and able to photograph each particle on the conveyor individually, as long as near-infrared (NIR)-spectra were detected from the particle (as NIR information is used by the machine for defining objects). Also, it must be noted that an individual particle was defined as a cluster of material-pixels, surrounded by a defined background, which was the conveyor belt. Therefore, overlapping or touching particles were detected as one object in an image and sorted out manually in the image-processing step. Images from objects smaller than 40 mm were taken with an ordinary digital camera. This size threshold was chosen because the small objects would have been difficult to handle in the experimental setup of the sensor-based sorting machine, because of the large-scale equipment with several transfer points on conveyor belts where the objects might have been lost during the process. The images were saved in *.jpg format, where each particle was photographed manually. Here, the resolution of the images for the software was originally set to 6,000x4,000 pixels, but later downsized and cropped to size in the software to roughly 400x500 pixels for faster processing. Despite the downsizing of the images, a resolution of 0.0511 mm^2/pixel was reached. In both cases of image acquisition, the cameras were mounted parallel to the surface where the objects were placed, and the position of the cameras was fixed to ensure a consistent distance between particle and camera lens. For manual image recording, this can be achieved using a tripod or similar. Additionally, LED-lights were used to ensure the high visibility of the particles and a well-lit area for the recording of the images. High contrast between the background and the particle itself helps to obtain images with sufficient quality for the further processing steps. In this case, all images were taken using a black background, which leads to the discarding of dark objects in the investigation, which were also not recorded by the sorter due to missing NIR-signal of black particles (Gundupalli et al., 2017). Due to the use of different methods for image acquisition, several steps were considered to ensure comparability between all images. Firstly, two (one for each image acquisition method) individual scaling factors – which allow the transformation between the pixel and metric measurements – were calculated as well as a distortion factor to remove the elongation which was detected on the images taken from the RGB-camera in the sorting machine due to the used combination of conveyor belt speed and imaging frequency of the line-camera. The scaling and distortion factors were calculated from images of objects with

known measurements and stored in variables in the software for later access.

The software evaluates calculated measurements of objects in binary images, which were converted from the RGB-photos. In MATLAB® this was achieved through transforming the original image (with the distortion factor considered, if required by the method of image acquisition) over a greyscale image (function "rgb2gray") to a binary data set (function "imbinarize"), where background information was indicated with zero (black) and pixels of the particle with one (white). The individual steps are shown in Fig. 2. Additionally, Table 1 gives an overview of all the used functions with a short description as well as additional information regarding in- and output data of each function. To receive the best results, several steps for image improvement were applied. These involve transforming the image to a negative image or enhancing the colours, where all entered parameters must be matched with the colour of the chosen background and/or particle. In MATLAB® the functions "imcomplement", "imreducehaze" and "imadjust" were used for these steps. Latter considered a "gamma factor" of 0.2, for the transformation to a binary image a global threshold according to Otsus Method (Otsu, 1979) was calculated (function "graythresh"), to achieve the best results.

Due to the image processing, small holes were detected in the binary data of the objects. In this context, holes are identified as regions indicated as background within the particle. To ignore falsely identified holes, all pixels of holes with a size smaller than 1% of the total image size were relabelled in the binary image from background to particle. Holes bigger than the chosen threshold were ignored in this step. Mechanical stressing from the sorting process of the material led to the separation of dust and fine particles, which were detected in the images. To ignore these objects in the evaluation only the biggest region of connected pixels was detected as the particle and displayed in the final binary image. The mentioned steps for identification and relabelling of holes as well as identifying the biggest region were achieved with the functions "bwareaopen", "imfill" and "regionprops".

To assure that the software only considers images where the binary data represents an object correctly (no overlapping, no cropped images due to missing contrast or NIR data) a manual check was necessary, where the best of three options was chosen. The first option considered the steps for image improvement as well as the gamma factor, which is stated in the text above. In the second option, the methods for colour improvement and the gamma factor are not applied. If the binary image was wrongly transformed, the third option allowed to not consider the image in the following analysis. The model here works with images of individual particles only, but research shows that the extraction of individual particles (no touching/overlapping) from one image is possible (Dunnu et al., 2006). MATLAB® allows this with the function "regionprops".

2.3. Calculation of particle descriptors

With the correctly displayed binary images available the next step was to calculate individual particle descriptors based on the binary data. To calculate the area of the particles, simply the number of pixels of the projected area was counted. The perimeter of the object was calculated with the option "bwboundaries". Due to the high image quality, the edge of the particles was displayed comparatively unevenly, and the length of the perimeter would have been wrongly calculated for the use here. To even out the irregular edge of the objects the perimeter was formed by a surrounding polygon with a shrink factor of 0.5 (see Fig. 3, left), which was achieved by applying the MATLAB® function "boundary". This resulted in the overall best approximation of the particle outline. It must be pointed out as well, that especially for objects with a low

Fig. 2. RGB image (left), greyscale image (middle), and binary image (right) of singlified waste particle.

number of pixels (small objects and/or low-resolution images), the squared shape of the pixels (resulting in a stair-like boundary) may affect the calculation of parameters. Since in this work, all relevant objects were far over a few hundred pixels in size (A_{Part}), this was not considered here. The Feret diameters were calculated with software-codes from (Eddins, 2018a, 2018b, 2017) (function "feretProperties"). This work considers the maximal as well as the minimal Feret diameter (see Fig. 3 (left)) because they are also successfully used in other works for particle description (Dražić et al., 2016; Olson, 2011; Podczeck, 1997).

The aim of using bounding shapes was to describe irregularly shaped particles more simply. In the work the following forms were determined: Smallest circumscribing rectangle (bounding box) of the particle, smallest circumscribing circle of the particle (bounding circle), largest circle in the circumscribing polygon (inscribed circle of the polygon) as well as the maximum inscribing circle of the particle (incircle) and the smallest circumscribing triangle of the particle. Of all the mentioned shapes, the dimensions that are needed to describe the shape (edge lengths, radius) were calculated and used to determine the circumference and area of the shape. Later these values were used to compare the measurements of the particle with the assigned bounding shapes. Fig. 3 (right) shows an example of some of the mentioned bounding shapes. In the software, functions from D'Errico (2014) ("minboundrect", "minboundcircle", "incircle", "minboundtri") and Birdal (2011) ("max_inscribed_circle") were used for the calculation.

Shape factors sf_i are dimensionless factors and show the ratio between different particle descriptive areas. Here, the following sf are considered: bounding box, bounding circle, bounding triangle, inner circle polygon and incircle. Additionally, the circularity was considered as a shape factor as well, which explains the difference of the particle from a circle. This factor was defined through Eq. (1) according to Hentschel and Page (2003), where A_{Part} is the projected area of the particle and P_{Part} is the perimeter of the circumscribing polygon, and was defined in a way, so that it becomes 1 for a circle.

$$Circularity = \frac{4 * \pi * A_{Part}}{P_{Part}^2} \tag{1}$$

The descriptors with additional information regarding particle size and material were calculated for each image and stored in table form, resulting in a data table containing the information regarding each particle in one row and the values for the individual descriptors in columns. All measurements were calculated in the unit pixel and were transformed in millimetre measurements. Here, the previously evaluated scale factor for the respective particle size class was considered in addition to the dimension of the unit of the descriptor (two dimensions for e. g. areas, one dimension for e.g. lengths, diameters and zero dimensions for shape fac-

tors). In the whole software the material fractions are examined separately, which leads to five individual data tables. These were used as input for the regression models, where one for each material class was considered.

2.4. Regression model

Prior investigations from this work show that the prediction performance of the particle sizes based on only one parameter performs poorly, because of the broad scattering of the values and collinearity between the variables (Kandlbauer, 2020). Therefore, the regression was performed with PLS1, where all the information of particle descriptors was used as predictor variables, and the information regarding particle size class was used as the one-dimensional response variable. The PLS method was chosen because it offers solutions for highly collinear or linearly dependent predictors (de Jong, 1993), which were determined in this case. The basic concept of PLS is stated with the following explanations and formulae from de Jong (1993), Massart et al. (1998) and Vandeginste et al. (1998a, 1998b).

The idea of PLS is to find a regression model for multivariate data for predicting purposes of new data by a small number of relevant factors. Here, the matrix of explanatory variables X is decomposed in a set of orthogonal factors which are used for fitting the matrix of response variables Y. In PLS latent variables (cannot be measured directly) are built that account for as much of the manifest variable (can be measured directly) variation as possible and are represented as linear combinations of the manifest variables. The final regression model is in the linear form of $Y = XB$, with B being regression coefficients which are calculated over the PLS factors between the explanatory and the response variables. The general model that underlies the PLS is given in Matrix form in equation (2) and (3) from Vandeginste et al. (1998b) as

$$X = TP^T + E \tag{2}$$

$$Y = TC^T + F \tag{3}$$

where X is an $n \times m$ matrix of predictors, Y is an $n \times p$ matrix of responses. T is an $n \times l$ matrix containing projections of X (*X-score*, *component* or *factor* matrix). P is an $m \times l$ orthogonal *loading* matrix, C is an $m \times l$ matrix and contains the regression coefficients which relate T to the variables in Y. The residuals are collected in the matrices E and F which are the error terms. PLS allows a variable reduction where the available variables are combined through linear combinations to fewer ones with more significance. Additionally, the method allows finding certain variables, that are essential for the description of the data and identify the ones that don't provide additional information. This can be done by interpreting the values of the weights w, which can be identified through the scores T.

Table 1

Overview of the used functions in the software with given source references. Note: if no information regarding In- and Output data is given in the Table, the function was automatically called from another function in the process.

Name of function	Description	Input	Output	Source reference
antipodalPairs	called by the function "feretProperties"			Eddins (2017)
boundary	boundary of a set of points	coordinates of the perimeter	point indices representing a single conforming 2-D boundary around the given points	MathWorks (2020a)
bwareaopen	remove small objects from binary image	binary image	binary image without small objects	MathWorks (2020b)
bwboundaries	traces the exterior boundaries of objects, as well as boundaries of holes inside these objects, in the binary image	binary image	pixel locations for boundaries	MathWorks (2020c)
feretDiameter	called by the function "feretProperties"			Eddins (2018a)
feretProperties	calculation of Feret diameters	binary image of (perimeter of) object	tables with edge lengths, endpoint coordinates, position (angles) of min. and max. Feret diameter	Eddins (2018a)
graythresh	global image threshold using Otsu's method	greyscale image	global threshold	MathWorks (2020d)
imadjust	adjust image intensity values or colormap	greyscale image	greyscale image with adjusted colour values	MathWorks (2020e)
imbinarize	binarize 2D greyscale image	greyscale image	binary image	MathWorks (2020f)
imcomplement	complement image	image data	complement image	MathWorks (2020g)
imfill	fill image regions and holes	binary image	filled binary image	MathWorks (2020h)
imreducehaze	reduce atmospheric haze	colour or greyscale image	dehazed image	MathWorks (2020i)
incircle	compute the maximal inner circle of the polygonal convex hull of a set of points in the plane	coordinates of the perimeter	radius and centre coordinates of the circle	D'Errico (2014)
inpoly	called by the function "max_inscribed_circle"			Birdal (2011)
max_inscribed_circle	compute the centre coordinates and radius of the maximum inscribed circle of a given object	Binary image of the particle outline	radius and centre coordinates of the circle	Birdal (2011)
maxFeretDiameter	called by the function "feretProperties"			Eddins (2018a)
minAreaBoundingBox	called by the function "feretProperties"			Eddins (2018a)
minboundcircle	compute the minimum radius of the enclosing circle of a set of points in the plane	coordinates of the perimeter	radius and centre coordinates of the circle	D'Errico (2014)
minboundrect	compute the minimal bounding rectangle of points in the plane	coordinates of the perimeter	coordinates that define the rectangle, area and perimeter of the rectangle	D'Errico (2014)
minboundtri	compute the minimum area bounding triangle of points in the plane	coordinates of the perimeter	coordinates that define the triangle	D'Errico (2014)
minFeretDiameter	called by the function "feretProperties"			Eddins (2018a)
pixelHull	called by the function "feretProperties"			Eddins (2018c)
plsregress	plsregress(X,Y,ncomp) computes a partial least-squares (PLS) regression of Y on X, using ncomp PLS components	predictor and response variable, amount of PLS components	predictor and response loadings and scores, PLS regression coefficients	MathWorks (2020j)
regionprops	measure properties of image regions	binary image	measurements for the set of defined properties for each 8-connected component in the binary image	MathWorks (2020k)
rgb2gray	convert RGB image or colourmap to	truecolour RGB image or colourmap	grayscale image	MathWorks (2020l)
signedTriangleArea	called by the function "feretProperties"			Eddins (2018a)
triangleHeight	called by the function "feretProperties"			Eddins (2018a)
zscore	returns the z-score for each element such that columns are centred to have mean 0 and scaled to have standard deviation 1	data with non-standardized values	data with standardized values	MathWorks (2020m)

To evaluate the models properly the available data tables were split into two separate groups, one containing 90% of the data from each particle size class, which was randomly picked. This data is later called the calibration data. The remaining data is used to test the quality of the developed model and is therefore called test data. In MATLAB® the function "plsregress" was applied, which is based on the algorithm from de Jong (1993). Additionally, the MATLAB® function "zscore" was used to standardize the variables in a way

that each column in the data set had a mean of 0 and a standard deviation of 1. A list of the used predictor variables and their explanation is given in Table 2. Due to the chosen methods of material preparation and image acquisition an unequal amount of images in each material-particle size class was available. To consider this fact in the regression model and avoid the more precise modelling of particle size classes with a higher number of images a way to compensate this factor was necessary. This was done by adding a
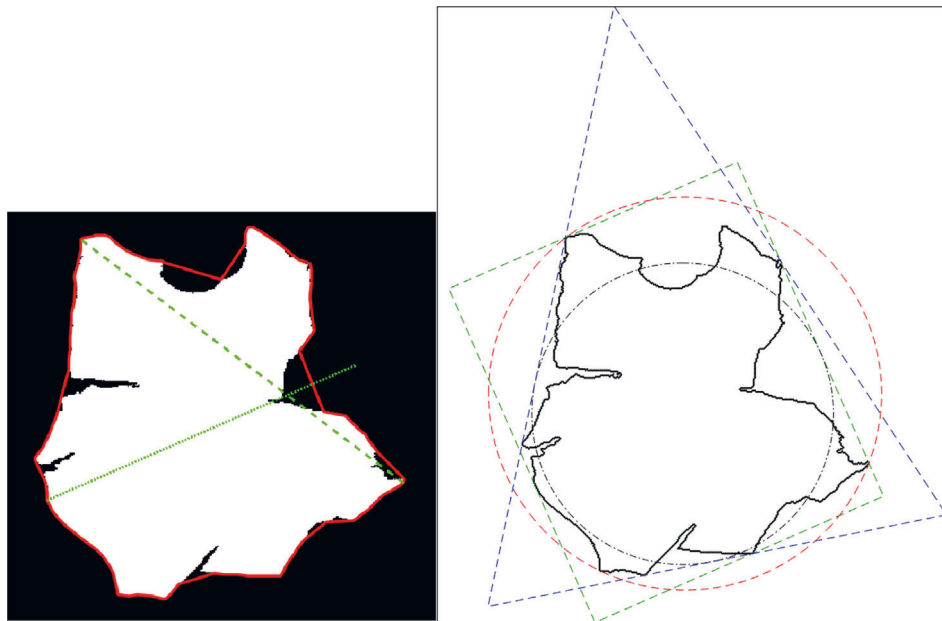
Fig. 3. Left: Particle with polygon perimeter (solid line), max. Feret diameter (dashed line) and min. Feret diameter (dotted line); right: Particle outline with exemplary bounding shapes (bounding box, bounding triangle, bounding circle (dashed line), inscribed circle of the circumscribing polygon (dash-dotted line)).

Table 2
Overview of the considered variables (descriptors) and their explanation.

Descriptor	Definition
Area particle (A_{Part})	projected area of the object
Perimeter (P_{Part})	length of the perimeter of the polygon of the projected area
Area bounding box	area of the smallest rectangle, which surrounds the object
Area bounding triangle	area of the smallest triangle, which surrounds the object
Shape factor[a]	ratio between the area of the object and area of the respective (bounding) shape
Circularity	explains difference from a circle and is defined through Eq. (1)
Diameter incircle	diameter of the largest inscribed circle of the particle
Feret diameter (maximal, minimal)	maximal/minimal distance between two parallel tangents, which fully enclose the object (principle of a calliper)

[a] Bounding box, bounding circle, inscribed circle of polygon, incircle, bounding triangle.

weighting factor, which is defined as the inverse value of the number of available images in each material-particle size class, respectively. Since the PLS is a method that allows the reduction of dimensions and eliminates noise in the data, the number of considered dimensions was chosen to be four. This was evaluated as the best option for the regression models, without under- and overfitting the data based on k-folds cross-validation with k being chosen to be 10, where at least 94% of the explained variance in Y was considered in the models.

The result of the models were regression coefficients for the predictor variables based on the calibration data, which were later applied to the test data sets to predict the particle size. Based on the fact that in this work five different particle size classes are present, the classification for the values to the particle size classes was implemented by considering a limit value for each particle size class. This limit value was calculated as the mean value between the predicted values of the 80th percentile of the lower particle size class and the 20th percentile of the higher particle size class (e. g. the limit value to divide particle size class 40–60 mm from 60–80 mm was the mean from the 80th percentile of the predicted values of the 40–60 mm particles and 20th percentile of the predicted values of the 60–80 mm particles).

3. Results and discussion

In this investigation in total over 11,000 particles were characterised with the chosen descriptors listed in Table 2. The breakdown according to particle size class and material fraction are shown in Table 3. Here, the values show the number of particles that were used for the regression models due to the manual selection of the binary images. The values in brackets show the share from the total amount of images that were taken from the cameras that is represented through the number of characterised particles. It can be detected that less than half of the amount of the available pictures in total were used in the regression. The following different factors explain the issue for particles ≥40 mm. Firstly, especially large objects were cropped and not fully displayed on one image, but on two or even more. This was also detected on particles with a very small expansion in one dimension, which often led to the detection of two or more individual objects. Additionally, images were often cropped when objects consisted of several different materials (composite materials) and the material was partly not reflecting NIR-signals, and therefore not detected by the sensor. Further, the surface condition was an important factor for image quality from particle size classes. If the particle was too dark in colour, see-through, too shiny, or had a reflective surface the image was often not properly transformed into a binary image and was not suitable for the regression model. These factors explain the significantly lower number of particles in the residual fraction when compared to the other materials, where e.g. glass, metal-coated plastics, composite materials and dark textile fibres and fabrics accounted for a high number of particles. Due to the composition and structure of the lightweight and the residual fraction (fluffy material that was tangled up with other particles) no photos were taken for these two material fractions in the particle size 10–20 mm. Also, it should be noted that the metal fraction was not considered in the evaluation. Here, a low number of objects was initially found in the samples. A significant amount

Table 3
Number and share of characterised particles regarding particle size class and material fraction. The numbers in brackets give the share of successfully analysed particles for each material-particle-size fraction.

	>80 mm		60–80 mm		40–60 mm		20–40 mm		10–20 mm		Total	
Plastics	1,422	(0.28)	748	(0.46)	998	(0.45)	409	(0.61)	241	(0.56)	3,818	(0.38)
Wood	495	(0.58)	452	(0.66)	409	(0.59)	350	(0.62)	225	(0.49)	1,931	(0.59)
Lightweight fraction	204	(0.19)	231	(0.36)	1,113	(0.45)	291	(0.57)	–	–	1,839	(0.39)
Paper-cardboard	2,061	(0.44)	266	(0.60)	232	(0.66)	363	(0.64)	119	(0.52)	3,041	(0.49)
Residual	61	(0.19)	70	(0.30)	133	(0.24)	237	(0.40)	–	–	501	(0.30)
Total	4,243	(0.35)	1,767	(0.49)	2,885	(0.46)	1,650	(0.57)	585	(0.52)	11,130	(0.42)

of these metallic objects were wires, which were often not correctly displayed on the images due to the one-dimensional appearance, flat metallic objects were often too shiny. Therefore, only an exceedingly small number of usable image files (less than 40 images in all particle size classes combined) was available and forced the fraction to be taken out from the investigation.

The given regression coefficients from the regression models were applied to the test data sets for each material fraction. The results were values that were assigned to a particle size class based on the classification method explained in the previous chapter. Because the size of the test data sets was dependent on the number of available images some material-particle size specific fractions held less than ten images. To evaluate the regression models in a way that even a small number of available objects gives meaningful information regarding the accuracy of the models the results presented here are calculated as a mean of 1,000 individual runs. In more detail, this means that 1,000 calibration data sets with data from randomly selected particles were used to generate the regression models. Each of the models was then tested by applying the resulting regression coefficients to the respective test data set and calculate the falsely and correctly assigned particles for each size class. These values can be seen as instant results from the regression models and are used for the prediction of the particle size. The mean results from 1,000 runs for the individual materials are presented in Table 4 as a significant result to show the accuracy of the method. Overall, the models for the materials wood and residual fraction show the best results regarding the correct

classification of the particles. However, the fraction paper-cardboard shows a very low correct classification of particles between the sizes 20–60 mm. To better evaluate the effects of the wrongly classified objects, the results from Table 4 are applied on particle size distributions. As no information regarding the individual particle weights was examined during this investigation, for this first approach it was assumed that each object in the same material-particle size class has the same weight. Two different options were investigated here.

Firstly, a uniformly distributed (UD) composition of the individual particle size classes within the material fractions was considered. Since five different particle classes were examined, each class was considered with a 20% mass fraction. For the material groups lightweight and residual fraction no images in the particle size class 10–20 mm were available, which is why these could not be included in the regression. In these cases, the particle size classes were each taken into account with 25% of the mass fraction.

Additionally, a realistic distribution (RD) of particle size classes within individual material fractions in mixed solid commercial waste according to Khodier et al. (2020) was considered. This distribution was chosen because it deals with the same input material, similar individual material fractions and similar particle size classes. The individual fractions are split up in more detailed material classes as in the present paper, but their combination allows the representation of the investigated fractions in this contribution. For this, the individual fractions "paper" and "cardboard" were summed up for the representation of the paper fraction. For

Table 4
Mean accuracy of the PLS-classification for 1,000 individual randomly picked test data sets in shares for the individual material fractions (wo: wood, 3D: 3D plastics, pa: paper&cardboard, 2D: lightweight fraction, re: residual fraction). Bold numbers represent the share for correctly classified objects.

	to 10–20 mm [%]	to 20–40 mm [%]	to 40–60 mm [%]	to 60–80 mm [%]	to >80 mm [%]	absolute number of particles [–]
wo from 10–20 mm	**100**	0	0	0	0	22
wo from 20–40 mm	0	**99**	1	0	0	35
wo from 40–60 mm	0	1	**88**	11	0	41
wo from 60–80 mm	0	0	12	**71**	17	45
wo from >80 mm	0	0	2	16	**82**	49
3d from 10–20 mm	**100**	0	0	0	0	24
3d from 20–40 mm	0	**100**	0	0	0	41
3d from 40–60 mm	0	0	**57**	39	4	100
3d from 60–80 mm	0	0	38	**51**	11	75
3d from >80 mm	0	0	1	12	**87**	142
pa from 10–20 mm	**100**	0	0	0	0	12
pa from 20–40 mm	0	**18**	38	44	0	36
pa from 40–60 mm	0	72	**15**	13	0	23
pa from 60–80 mm	0	4	8	**88**	0	27
pa from >80 mm	0	0	0	0	**100**	206
2d from 20–40 mm	0	**72**	27	1	0	29
2d from 40–60 mm	0	29	**65**	6	0	111
2d from 60–80 mm	0	0	7	**65**	28	23
2d from >80 mm	0	0	2	25	**73**	20
re from 20–40 mm	0	**100**	0	0	0	24
re from 40–60 mm	0	0	**100**	0	0	13
re from 60–80 mm	0	0	0	**96**	4	7
re from >80 mm	0	0	0	6	**94**	6

Table 5
Material-specific error (in %) between the mass share of the predicted regression model and the mass share of the considered distribution (UD, RD) for each particle size class, as well as the mean error per particle size class (wo: wood, pa: paper&cardboard, 3D: 3D plastics, re: residual fraction, 2D: lightweight fraction, UD: uniformly distributed, RD: realistic distribution).

particle size [mm]	wo UD	wo RD	pa UD	pa RD	3D UD	3D RD	re UD	re RD	2D UD	2D RD
10–20	0	–	0	–	0	–	0	–	0	–
20–40	0	0	−1	1	0	0	0	0	0	1
40–60	1	1	−8	−9	−1	1	0	0	0	1
60–80	−1	−1	9	8	0	2	0	2	0	14
>80	0	0	0	0	1	−3	0	−2	0	−16
Error per particle size class	0.2	0.25	1.8	2.25	0.2	0.75	0	0.5	0	4

the residual fraction the values from the "inert material" and "textiles" were added to the "residual fraction". The used values for the further calculations are presented in the supplementary material (see Table S1 in Appendix). Since in that study no data according to material-specific composition in fractions smaller 20 mm was investigated, the particle size class 10–20 mm was not considered in the results presented here.

The aforementioned assumption regarding particle masses combined with the two options of particle size distributions allows to constitute the impacts of the regression models and compare the original particle size distribution with the one predicted through the respective model. The extent of incorrectly assigned particles on the particle size distribution is shown, considering the partially compensating effect of wrong classifications. The results of the according hypothetical sensor-based screening analysis in form of the percentual error of the mass fractions between the predicted values and the original particle size distribution is shown in Table 5. Furthermore, the errors of the models can be given by calculating the mean error in terms of mass share for all particle size classes, which are also presented in Table 5. Here, it is shown that the approach of uniformly distributed (UD) material regarding particle size classes shows overall better results than considering the realistic particle size distribution (RD). Nevertheless, the accuracy of the predicted particle size distribution for the fractions wood, plastics (3D) and residuals was at least 99% for each particle size class respectively. It is also shown that the error between UD and RD for each material is in the same range, except for lightweight material. Here, the difference in error can be interpreted by the difference between the particle size shares in the distributions, where RD considers approx. 70% of particles > 80 mm, while UD just 25%.

4. Conclusions

The presented methodology, which predicts the particle size based on parameters from 2D images, shows promising results for measuring the particle size distribution of the investigated material fractions. Considering a realistic particle size distribution with the approach of identical particle weights in each particle size class the regression model led to a correct prediction of at least 99% the individual particle size classes for wood, plastics (3D) and residual fraction. The fractions paper-cardboard and lightweight materials showed significant errors but were still over 96% accurate for each particle size class respectively. Nevertheless, the following points should be noted as a limitation of the stated method.

For the recording of images by the RGB sensor of the sensor-based sorting machine a detected NIR-signal was crucial. This factor mainly caused dark (especially black and grey) objects not to be considered in the evaluation. Additionally, certainly shaped objects (one-dimensional) were recorded on multiple separate images and therefore not useable in the investigation. The chosen way of image processing in the software requires a manual check of the binary images. This is not feasible for a practical application and

could not be implemented due to the real-time request in a plant. Hence, a software extension is necessary, by which this this control can be carried out automatically to eliminate the manual effort.

Another point of criticism is that large objects were often cropped on the images due to the limiting size of the images. This led to missing particle information in bigger particle size classes. The fine fraction below 10 mm was not considered in the investigation at all but is crucial information when evaluating particle size distributions.

The lightweight fraction was classified by the air classifier and consisted mostly of plastic foils. Additionally, the whole material stream was checked by hand as well to avoid wrongly classified objects (e. g. foils wrapped up with wires). In this paper, an approach to evaluate the material-specific particle size class from the whole (mixed) material stream was investigated. Especially the differentiation between the materials lightweight fraction and 3D plastics based only on the information from a NIR-sorting machine is very limited in a real plant, since both fractions include objects from the same materials (Möllnitz et al., 2020). Also, an efficient singlification of the material stream on a conveyor for sufficient detection of particle shapes is not always possible due to limited space and upkeeping of the mandatory throughput of the plants. Here, a bypass for smaller quantities could be used as a technical solution for material analysis. Additionally, the image recording must be tested when the particle size classes are not separately recorded by the cameras. Here, especially large foils >80 mm would probably cover up smaller particles on the conveyor belt, so that a screening step might be necessary. Nevertheless, although further research is needed to develop a Smart Waste Factory, the presented approach shows high potential to be used as an automated method to measure particle size distributions of solid, mixed waste streams to process large amounts of data.

Besides, particle weights must be examined and combined with the models to evaluate the effect of the wrongly classified particles for more realistic information regarding particle size distribution. I. e. waste management will become particle-, sensor- and data-based-management and information from every single particle will become increasingly relevant in the future, especially because of higher recycling rates set up by the EU Circular Economy Package.

Further, approaches regarding machine learning could be investigated to evaluate if additional information about particle shapes can be extracted based on the chosen particle descriptors, or other information in the images. The images (true colour or greyscale) can also be processed over Convolutional Neural Networks, which can detect characteristic shapes and edges of objects and may lead to better identification of shapes.

Ultimately, it must be mentioned that the results are based on material that is classified by a drum screen. If the method should be applied on other screen types or drum screens with different screen perforations, the efficiency of the respective screen on material fractions and particle size classes must be investigated separately but could give information about particle sizes when combined with the presented method here.

Funding

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wasman.2020.11.003.

References

Birdal, T., 2011. Maximum Inscribed Circle using Distance Transform. https://de.mathworks.com/matlabcentral/fileexchange/30805-maximum-inscribed-circle-using-distance-transform (accessed 27 July 2020).

D'Errico, J., 2014. A suite of minimal bounding objects. https://de.mathworks.com/matlabcentral/fileexchange/34767-a-suite-of-minimal-bounding-objects?focused=3820650&tab=function (accessed 23 April 2020).

Dražić, S., Sladoje, N., Lindblad, J., 2016. Estimation of Feret's diameter from pixel coverage representation of a shape. Pattern Recogn. Lett. 80, 37–45. https://doi.org/10.1016/J.PATREC.2016.04.021.

Dunnu, G., Hilber, T., Schnell, U., 2006. Advanced size measurements and aerodynamic classification of solid recovered fuel particles. Energy Fuels 20 (4), 1685–1690. https://doi.org/10.1021/ef0600457.

Eddins, S., 2017. Feret Diameters and Antipodal Vertices. https://blogs.mathworks.com/steve/2017/10/24/feret-diameters-and-antipodal-vertices/ (accessed 23 April 2020).

Eddins, S., 2018a. Feret Properties – Wrapping Up. https://blogs.mathworks.com/steve/2018/04/17/feret-properties-wrapping-up/ (accessed 23 April 2020).

Eddins, S., 2018b. Minimum Feret Diameter. https://blogs.mathworks.com/steve/2018/02/20/minimum-feret-diameter/ (accessed 23 April 2020).

Eddins, S., 2018c. What ist the Shape of a Pixel? https://blogs.mathworks.com/steve/2018/03/16/what-is-the-shape-of-a-pixel/ (accessed 23 April 2020).

European Commission, 2018. Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste.

Flamme, S., Hams, S., Kölking, M., 2018. ARGOS – Entwicklung eines Multisensor-Systems zur Echtzeitanalyse von metallreichen Aufbereitungsprodukten [Development of a multisensor-system for real-time analysis of metal-rich processing products], in: Vorträge-Konferenzband zur 14. Recy & DepoTech-Konferenz. Recy & DepoTech-Konferenz, Leoben, pp. 119–124.

Gundupalli, S.P., Hait, S., Thakur, A., 2017. A review on automated sorting of source-separated municipal solid waste for recycling. Waste Manag. 60, 56–74. https://doi.org/10.1016/j.wasman.2016.09.015.

Hentschel, M.L., Page, N.W., 2003. Selection of descriptors for particle shape characterization. Part. Part. Syst. Char. 20 (1), 25–38. https://doi.org/10.1002/ppsc.200390002.

de Jong, S., 1993. SIMPLS: an alternative approach to partial least squares regression. Chemom. Intell. Lab. Syst. 18 (3), 251–263. https://doi.org/10.1016/0169-7439(93)85002-X.

Kandlbauer, L., 2020. Sensorische Messung der Trommelsieb-Korngrößenverteilung von gemischten Gewerbeabfällen [Sensor-based measurement of particle size distribution of mixed commercial waste]. Masterarbeit [Master thesis]. Leoben.

Khodier, K., Curtis, A., Sarc, R., Lehner, M., O'Leary, P., Pomberger, R., 2019. Smart solid waste processing plant vision and pathway. Conference Contribution ISWA World Congress 2019.

Khodier, K., Feyerer, C., Möllnitz, S., Curtis, A., Sarc, R., (unpublished). Efficient derivation of significant results from mechanical processing experiments with mixed solid waste. in submission.

Khodier, K., Viczek, S.A., Curtis, A., Aldrian, A., O'Leary, P., Lehner, M., Sarc, R., 2020. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis. Int. J. Environ. Sci. Technol. 17 (2), 959–972. https://doi.org/10.1007/s13762-019-02526-w.

Kontny, M., 2017. Machine vision methods for estimation of size distribution of aggregate transported on conveyor belts. Vibroengineering PROCEDIA 13, 296–300 https://doi.org/10.21595/vp.2017.19151.

Krämer, P., 2017. Entwicklung von Berechnungsmodellen zur Ermittlung relevanter Einflussgrößen auf die Genauigkeit von Systemen zur nahinfrarotgestützten Echtzeitanalytik von Ersatzbrennstoffen. Shaker Verlag GmbH, Aachen.

Kuhl, F.P., Giardina, C.R., 1982. Elliptic Fourier features of a closed contour. Int. J. Image Graph. 18 (3), 236–258. https://doi.org/10.1016/0146-664X(82)90034-X.

Kuhn, M., Johnson, K., 2016. Applied Predictive Modeling. Springer-Verlag, New York, p. 600.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J., 1998. Principal components, in: Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J. (Eds.), Handbook of Chemometrics and Qualimetrics: Part A, 20th ed. Elsevier, pp. 519–556.

MathWorks, 2020a. boundary. https://de.mathworks.com/help/matlab/ref/boundary.html (accessed 26 June 2020).

MathWorks, 2020b. bwareaopen. https://de.mathworks.com/help/images/ref/bwareaopen.html (accessed 26 June 2020).

MathWorks, 2020c. bwboundaries. https://de.mathworks.com/help/images/ref/bwboundaries.html (accessed 26 June 2020).

MathWorks, 2020d. graythresh. https://de.mathworks.com/help/images/ref/graythresh.html (accessed 26 June 2020).

MathWorks, 2020e. imadjust. https://de.mathworks.com/help/images/ref/imadjust.html (accessed 26 June 2020).

MathWorks, 2020f. imbinarize. https://de.mathworks.com/help/images/ref/imbinarize.html (accessed 26 June 2020).

MathWorks, 2020g. imcomplement. https://de.mathworks.com/help/images/ref/imcomplement.html (accessed 26 June 2020).

MathWorks, 2020h. imfill. https://de.mathworks.com/help/images/ref/imfill.html (accessed 26 June 2020).

MathWorks, 2020i. imreducehaze. https://de.mathworks.com/help/images/ref/imreducehaze.html (accessed 26 June 2020).

MathWorks, 2020j. plsregress. https://de.mathworks.com/help/stats/plsregress.html (accessed 26 June 2020).

MathWorks, 2020k. regionprops. https://de.mathworks.com/help/images/ref/regionprops.html (accessed 26 June 2020).

MathWorks, 2020l. rgb2gray. https://de.mathworks.com/help/matlab/ref/rgb2gray.html (accessed 26 June 2020).

MathWorks, 2020m. zscore. https://de.mathworks.com/help/stats/zscore.html (accessed 26 June 2020).

Möllnitz, S., Khodier, K., Pomberger, R., Sarc, R., 2020. Grain size dependent distribution of different plastic types in coarse shredded mixed commercial and municipal waste. Waste Manag. 103, 388–398. https://doi.org/10.1016/j.wasman.2019.12.037.

Olson, E., 2011. Particle shape factors and their use in image analysis-Part 1: Theory. J. GXP Compl. 15–3, 85–96.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst., Man, Cybern. 9 (1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076.

Peddireddy, S., Longhurst, P.J., Wagland, S.T., 2015. Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste manag. 40, 9–13. https://doi.org/10.1016/j.wasman.2015.03.015.

Podczeck, F., 1997. A shape factor to assess the shape of particles using image analysis. Powder Technol. 93 (1), 47–53. https://doi.org/10.1016/S0032-5910(97)03257-9.

Pomberger, R., Sarc, R., Lorber, K.E., 2017. Dynamic visualisation of municipal waste management performance in the EU using Ternary Diagram method. Waste Manag. 61, 558–571. https://doi.org/10.1016/j.wasman.2017.01.018.

Sarc, R., Curtis, A., Kandlbauer, L., Khodier, K., Lorber, K.E., Pomberger, R., 2019. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management – a review. Waste Manag. 95, 476–492. https://doi.org/10.1016/j.wasman.2019.06.035.

Sarc, R., Lorber, K.E., 2013. Production, quality and quality assurance of Refuse Derived Fuels (RDFs). Waste Manag. 33 (9), 1825–1834. https://doi.org/10.1016/j.wasman.2013.05.004.

Sarc, R., Lorber, K.E., Pomberger, R., Rogetzer, M., Sipple, E.M., 2014. Design, quality, and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry. Waste Manag. Res. 32 (7), 565–585. https://doi.org/10.1177/0734242X14536462.

Sarc, R., Seidler, I.M., Kandlbauer, L., Lorber, K.E., Pomberger, R., 2019b. Design, quality and quality assurance of solid recovered fuels for the substitution of fossil feedstock in the cement industry – update 2019 734242X19862600 Waste Manag. Res. https://doi.org/10.1177/0734242X19862600.

Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J., 1998a. Multivariate calibration, in: Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J. (Eds.), Handbook of Chemometrics and Qualimetrics: Part B, vol. 20, 20th ed., Elsevier, pp. 349–381.

Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J., 1998b. Relations between measurement tables, in: Vandeginste, B.G.M., Massart, D.L., Buydens, L.M.C., Jong, S. de, Lewi, P.J., Smeyers-Verbeke, J.

(Eds.), Handbook of Chemometrics and Qualimetrics: Part B, vol. 20, 20th ed., Elsevier, pp. 307–347.

Yang, W.C., 2003. Handbook of Fluidization and Fluid-particle Systems. Marcel Dekker, New York, p. 861.

Zhang, Z., Yang, J., Su, X., Ding, L., 2013. Analysis of large particle sizes using a machine vision system. Physicochem. Probl. Miner Process 49, 397–405. https://doi.org/10.5277/ppmp130202.

Zimmermann, I., 1998. Teilchengrößenanalyse [Analysis of particle size], in: Zimmermann, I. (Ed.), Pharmazeutische Technologie [Pharmaceutical Technology], vol. 3. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 245–300.

Zlatev, M., 2005. Beitrag zur quantitativen Kornformcharakterisierung unter besonderer Berücksichtigung der digitalen Bildaufnahmetechnik [Contribution to quantitative characterization of the particle shape with special consideration of digital imaging technology]. Dissertation. Freiberg.

4   SUMMARY AND DISCUSSION

Here, each research question (see Section 2.3) is answered in detail and complemented with a summary of explanatory results and findings from own investigations.

4.1 Answer to research question 1

How can the processing products be representatively sampled in coarse-shredding experiments with mixed commercial waste, and how large are the remaining sampling errors?

In accordance with the theory of sampling, coarse-shredding products can be representatively sampled by forming a composite sample from multiple increments, taken from the stream, falling from the output conveyor belt during defined time intervals. Calculating the total sample mass, based on the Austrian standard ÖNORM S 2127, the reaming sampling errors for various particle size-material classes in a coarse-shredding experiment with mixed commercial waste were quantified through a so-called replication experiment (see Table 2). They are expressed as relative sampling variabilities (RSV): the standard deviations of the shares' classes divided by their estimated true values.

Table 2: Relative sampling variabilities (RSV) for various particle size classes and for the material classes metal (ME), wood (WO), paper (PA), cardboard (CB), 2D plastics (2D), 3D plastics (3D), inert materials including glass (IN), textiles (TX), and a residual fraction (RE) (publication I: Khodier et al., 2020)

Particle class [mm]	ME (%)	WO (%)	PA (%)	CB (%)	2D (%)	3D (%)	IN (%)	TX (%)	RE (%)	Sum (%)
0–5	-	-	-	-	-	-	-	-	12.3	12.3
5–10	-	-	-	-	-	-	-	-	12.3	12.3
10–20	-	-	-	-	-	-	-	-	10.4	10.4
20–40	41.4	17.7	24.3	39.3	18.4	17.1	19.7	29.3	22.7	11.6
40–60	47.3	21.5	16.8	25.6	14.2	8.7	37.2	43.4	16.4	8.8
60–80	39.4	23.3	22.9	18.1	17.7	10.0	49.9	30.4	8.9	8.1
80–100	62.0	34.7	38.0	14.2	19.3	17.5	210.7	43.4	17.2	7.7
100–200	74.0	47.7	69.0	21.6	28.9	35.2	131.8	40.9	40.0	10.9
200–400	153.0	230.9	203.9	126.2	38.3	39.8	-	42.9	52.2	28.8
Sum	16.4	18.3	10.5	15.0	16.6	12.1	31.2	26.6	3.6	0.0

RSV < 20%	20% ≤ RSV < 50%	RSV ≥ 50%

This research question is addressed in detail in **publication I**.

Pierre Gy's theory of sampling provides the theoretical basis for thoroughly understanding representative sampling. Its central requirement is accordance with the fundamental sampling principle: each particle must have the same probability of ending up in the final sample. According to the theory of sampling, there are intrinsic, and avoidable sampling errors. The latter – which are called incorrect sampling errors – are divided into three kinds of errors: increment delimitation errors, increment extraction errors, and incorrect processing errors.

Increment delimitation refers to the subdivision of the lot into segments of constant volume, building the potential increments for extraction. For coarse-shredding, the material leaves the shredder through an output conveyor belt. Hence, a practicable approach for increment de-limitation, which is in accordance with the fundamental sampling principle, is defining segments of time for sampling from the stream falling from the conveyor belt.

Increment extraction refers to correctly extracting the chosen delimited increments. In the replication experiment in this thesis, increment extraction was implemented by catching ma-terial from the falling stream with a sampling bucket covering the stream's whole cross-section during defined periods. Practice showed that this is a suitable but not perfect ap-proach: the manual procedure leads to variations in sampling duration. Hence, machine-assisted methods (as applied in publication III, for example) are preferable.

Incorrect processing errors are – more precisely – errors introduced after sampling and not during sampling itself. There are six kinds of such errors: contamination by foreign material, loss of material, alteration in chemical and alteration in physical composition, involuntary op-erator faults, and deliberate faults for manipulating results. These kinds of errors are not spe-cific to coarse-shredding but depend on the subsequent processing of the samples.

Besides the incorrect sampling errors, there are two kinds of so-called correct sampling er-rors, which are unavoidable due to the nature of sampling: the fundamental sampling error and the grouping and segregation error. The first is caused by the material's constitutional heterogeneity, which describes chemical and physical differences between fragments of the lot. While it affects the precision of the data, it does not harm representativeness. The fun-damental sampling error can only be reduced by either increasing the sample mass or changing the physical condition of the particles through comminution. For the case of so-called simple particles, it can be calculated with Equation (7) in publication I. Simple particles are such, of which 100% is assigned to a particular analyte, as is the case in particle size class analyses and waste sorting analyses.

The grouping and segregation error is caused by the distributional heterogeneity, which re-fers to the spatial distribution of different particles. It can be faced by mixing the lot or by in-creasing the number of increments that form a sample, and thereby the spatial coverage of the lot.

The equation for calculating the fundamental sampling error requires assumptions on the composition and average particle weights of the material and only covers one out of two cor-rect sampling errors. Consequently, in this thesis, the sample mass was defined based on the Austrian standard ÖNORM S 2127 (see Equation (11) in publication I), which incorpo-rates a linear dependence of the sample mass on the 95th percentile particle size. This linear-ity assures feasible sample masses in terms of practicability. Still, it neglects that a quadratic

to cubic correlation would be necessary to consider the particles' dimensionality, according to the theory of sampling. The performance of the established procedure, combined with the sample mass retrieved from the standard, was hence evaluated based on a replication experiment. In such a replication experiment, the whole procedure of sampling and analysis is performed in parallel at least ten times, to quantify the variation of the analyses' results.

The overall procedure for sampling, and screening and sorting analysis established in this thesis is shown in Fig. 3 in publication I. The resulting sampling errors (Table 2) range from 3.6% for the total content of a residual fraction to about 231% for the content of wood from the particle size class of 200–400 mm. They show reasonable errors with a maximum of 31.2% when investigating either the overall composition or the overall particle size distribution of the waste and get much worse when the level of detail is increased, analyzing particle size-material classes.

This observation is in accordance with the dependence of the detected sampling errors on the share of the analyte, as shown in Figure 3. The Figure also visualizes the trend of the fundamental sampling error over the analyte's share. Considering typical average particle weights – which are much lower than 1 kg for most particle size-material classes – it shows that for the chosen procedure significant errors beyond the fundamental sampling errors are present – most likely grouping and segregation errors.

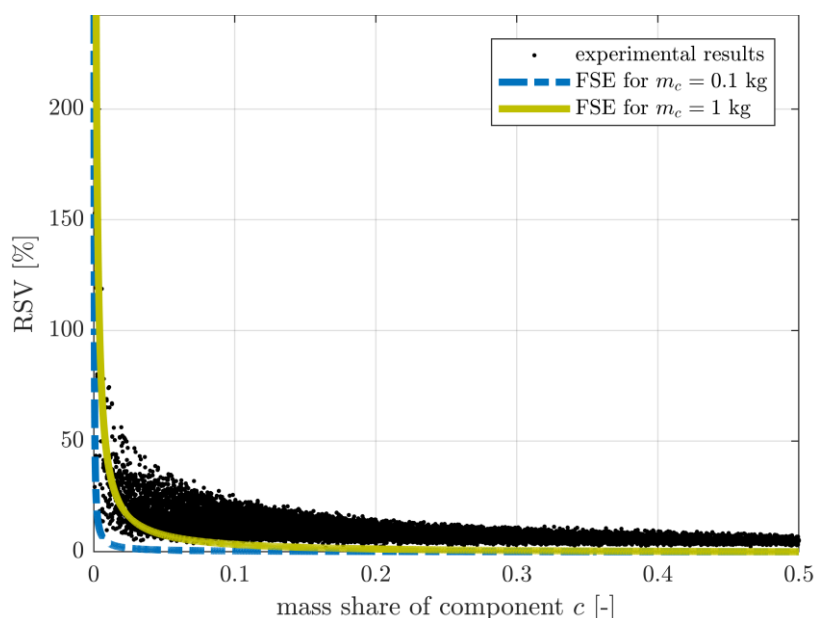Figure 3: Relative sampling variability (RSV) for the concentration of a component c versus its share: experimental results, and theoretical fundamental sampling errors (FSE) for two different average particle masses m_c of component c, a lot mass of 45,000 kg, a sample mass of 240 kg, and an average particle mass of 0.1 kg for other components (combined from Fig. 6 and Fig. 7 in publication I – Khodier et al., 2020)

4.2 Answer to research question 2

How can reliable conclusions about machine influences on waste shredding be drawn, despite sampling errors and inter-experimental differences in the waste?

Reliable conclusions about machine influences on waste shredding require distinguishing between real effects and apparent effects caused by the distortion introduced by the waste's heterogeneity. The real effects can be identified in an analysis of variance (ANOVA), based on polynomial regression models, calibrated with a sufficiently large amount of data from randomized, Design of Experiments-based investigations.

This research question is addressed in detail in **publication II**.

Sampling errors and inter-experimental differences contribute to the distortion and thereby introduced residual variance of the data that is not explainable by models on factor influences. In the presence of high residual errors, the detection of effects requires a sufficiently large amount of data so that real effects become significant in suitable statistical tests. Furthermore, a randomization of the order of the runs is desirable since it randomly spreads the influence of the waste's condition over different factor settings.

The method Design of Experiments allows choosing factor settings for experimental runs that maximize the efficiency of the experiment in terms of extractable information per run. In this thesis, a D-optimal design was selected from the various existing designs – the numerically built family of optimal designs allows the efficient combination of nominal (cutting tool geometry) and numerical (gap width and shaft rotation speed) factors and is flexible concerning the chosen underlying design model.

An ANOVA subsequently distinguishes between significant, real, and non-significant, apparent effects. Finally, the obtained model is validated by checking the data for outliers and analyzing the residuals on the normality of their distribution and non-desired trends regarding runs, model values, and factor settings.

In this thesis, a coarse-shredding experiment with mixed commercial waste and 32 runs of one hour was performed, investigating the influence of the radial gap width, the shaft rotation speed, and three cutting tool geometries on the throughput behavior and energy demand. It was performed and analyzed, using the described methods, applying a D-optimal design with a reduced cubic design equation (see Equation (1) in publication II). Significant models for the influence of these factors on the examined process parameters were successfully derived, with p-values <0.0001. They explain between 73% (hourly mass flow) and 81% (volume flow steadiness) of the variation in the experimental data. These results prove the capability of the method for drawing trustworthy conclusions on machine influences on waste shredding.

4.3 Answer to research question 3

How can particle size distributions be empirically modeled?

Particle size distributions can be empirically modeled by describing them as isometric log-ratio-transformed shares of particle size classes, applying multivariate multiple linear regression modeling, and identifying significant factors based on a multivariate analysis of variance (MANOVA).

This research question is addressed in detail in **publication III**.

Particle size distributions are a non-scalar property of particulate materials. Consequently, answering the research question concerns two issues: how to describe the particle size distribution and how to handle this kind of description as a dependent variable.

Common descriptive approaches consist of three classes, with distinct advantages and disadvantages: Lists of particles' sizes preserve detailed information on each particle. But this level of detail gets in the way of straightforward interpretation and does not summarize any trends in a suitable form for empirical modeling. Summary values, like the Sauter diameter, aggregate the information to a small number of values or even to a single scalar value. While this aggregation is favorable for modeling, it does not always preserve all relevant information on the size distribution of the particles. And analytical probability density functions provide detailed information on the distributions while describing them by a small number of momentums – but they are only applicable to particle size distributions produced by a small number of processes.

Hence, in this thesis, a distribution-independent approach that allows preserving a deliberate choice of information was searched for: a description through isometric log-ratios of a selection of particle size classes allows summarizing the distribution, maintaining exactly the relevant information for the specific case, while satisfying the mathematical constraints, introduced by the compositional nature of shares of particle size classes.

Modeling the isometric log-ratio-transformed shares of particle size classes can be performed by extending the methods presented in publication II to multivariate dependent variables: the result is a multivariate multiple linear regression model obtained from a MANOVA. Applying the proposed approach, a significant and validated model for the particle size distribution of mixed commercial waste, comminuted by coarse shredders (see the experiment, described in section 4.2), was found, confirming the suitability of the method. For three particle size classes >80 mm, 30–80 mm, and 0–30 mm, the model identifies the cutting tool geometry as the only significant factor, explaining 57% of the variance in the data.

4.4 Answer to research question 4

How big is the influence of the radial gap width, the shaft rotation speed, and the cutting tool geometry of an industry-scale coarse shredder on the throughput behavior, the energy demand, and the particle size distribution of shredded mixed commercial waste?

For the investigated machine (a Komptech Terminator 5000 SD) and within the examined design space, the maximum influences of changing the radial gap width, shaft rotation speed, or cutting tool geometry on the throughput, throughput steadiness, and specific energy demand at average settings of the corresponding other two factors are shown in Table 3.

Table 3: Maximum changes of the throughput (mass flow and volume flow), throughput steadiness, and energy demand caused by changes of the gap width, shaft rotation speed, and cutting tool geometries at average settings of the other two factors

	Gap width	Shaft rotation speed	Cutting tool geometry
Volume flow [m³/h]	38	43	95
Mass flow [t/h]	4.6	10.8	7.1
Volume flow steadiness [-]	0.02	0.08	0.22
Mass flow steadiness [-]	0.01	0.07	0.12
Specific energy demand [L/t]	0.45	0.45	1.43

Concerning the volume flow, the volume flow steadiness, and the mass flow steadiness, the influences of the gap width and the cutting tool geometry are interdependent. Consequently, individual considerations of the three cutting tools (F, XXF, and V) show that changes of the gap width influence the volume flow by up to 71 m³/h (XXF), the volume flow steadiness by up to 0.10 (V), and the mass flow steadiness by up to 0.10 as well (V and XXF). And when regarding the whole range of examined radial gap widths, the maximum effects of the choice of the cutting tool geometries are 129 m³/h for the volume flow (changing between F and V at maximum gap width), 0.38 for the volume flow steadiness (changing between XXF and V at minimum gap width), and 0.30 for the mass flow steadiness (changing between XXF and V at maximum gap width).

Regarding the particle size distribution, in terms of the examined size classes >80 mm, 30–80 mm, and 0–30 mm, it is only significantly affected by the cutting tool geometry,

with a maximum effect when switching between F and V (which changes the size class of 18% of the comminution product).

Publications II and III address this research question in detail.

The discussed effects refer to the expected model values, according to the empirical models, which were derived, using the methods addressed in sections 4.2 and 4.3. Their documented maximum magnitudes concern the examined design space, which covers the three cutting tool geometries F, XXF, and V, radial gap widths from 0% (0 mm) to 100% (33 mm for F, 35 mm for XXF, and 30 mm at the bottom of the teeth and 38 mm at the top of the teeth for V), and shaft rotation speeds of 60% to 100% (of a maximum of 31 rpm).

The trends of the detected significant average effects on the throughput behavior and energy demand are plotted in Figure 4. They refer to the average settings of the corresponding other factors, i.e., a gap width of 50%, a shaft rotation speed of 80%, and average observations of all three cutting tool geometries. The interactions between the radial gap width and the cutting tool geometry are visualized in Figure 5. Finally, Figure 6 shows the confidence and prediction regions of the composition of the three examined particle size classes, depending on the choice of the cutting tool geometry.

Figure 4 demonstrates that increasing the gap width leads to a linear increase of the volume flow, with strong effects for the F and XXF geometries and a non-significant (negative) influence for the V geometry (Figure 5). The mass flow also linearly increases with the gap width, while the specific energy demand decreases almost linearly. And the throughput steadiness is hardly affected in average of the three cutting tool geometries (Figure 4), and also at individual consideration of the F geometry (Figure 5), while the volume flow steadiness and mass flow steadiness show extrema at a gap width of around 50% for the other two geometries – a maximum for XXF and a minimum for V.

As for the shaft rotation speed, Figure 4 shows that raising it results in a linear increase of the volume flow and a quadratic change of the mass flow, with a maximum at about 84% of the maximum rotation speed – a discrepancy that is discussed in detail in publication II. The volume flow steadiness and mass flow steadiness decrease linearly, with increased shaft rotation speeds. And the quadratic dependence of the energy demands' logarithm on the shaft rotation speed results in an almost symmetrical curve with a minimum specific energy demand at around 80% of the maximum shaft rotation speed.

Concerning the cutting tool geometry, the F tool and XXF tools behave quite similarly in terms of the mass flow, specific energy demand, and throughput steadiness (Figure 4), with a slight dependence of the latter on the gap width (Figure 5). The volume flow is lower for the XXF geometry than for the F geometry (Figures 4 and 5), which is coherent with the finer material it produces, according to Figure 6. The most significant differences, though, are

those between the V geometry and the other two: it produces the finest material of the three (Figure 6), at the cost of higher specific energy demand, and lower but steadier throughput, in terms of volume and mass (Figures 4 and 5).
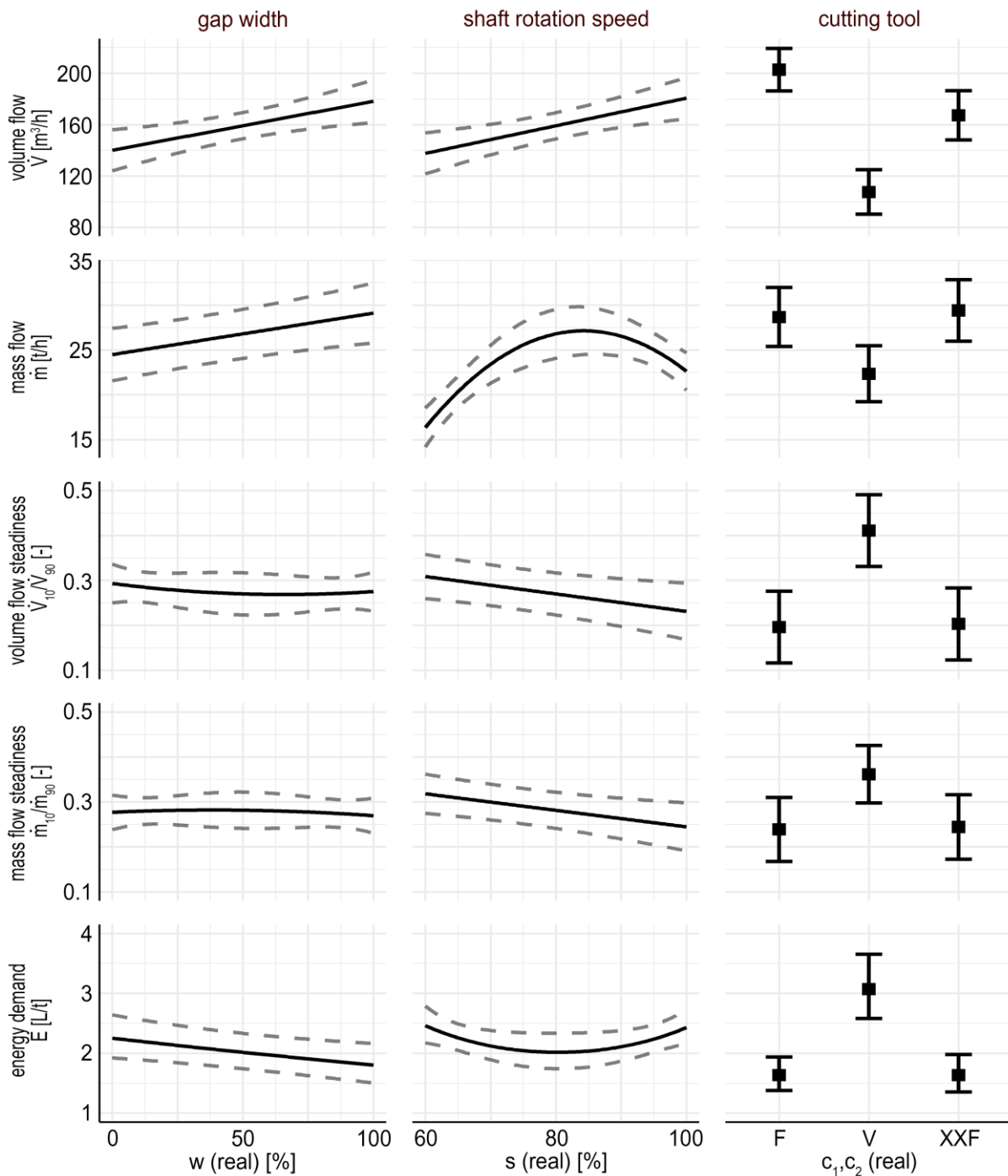
Figure 4: Effect plots with confidence bands for the influence of gap width (w), shaft rotation speed (s), and cutting tool geometry (c) on throughput, throughput steadiness, and specific energy demand, at the average setting of the respective other factors (redesigned version of Fig. 3 in publication II – Khodier et al., 2021 – with absolute volume flows and mass flows)
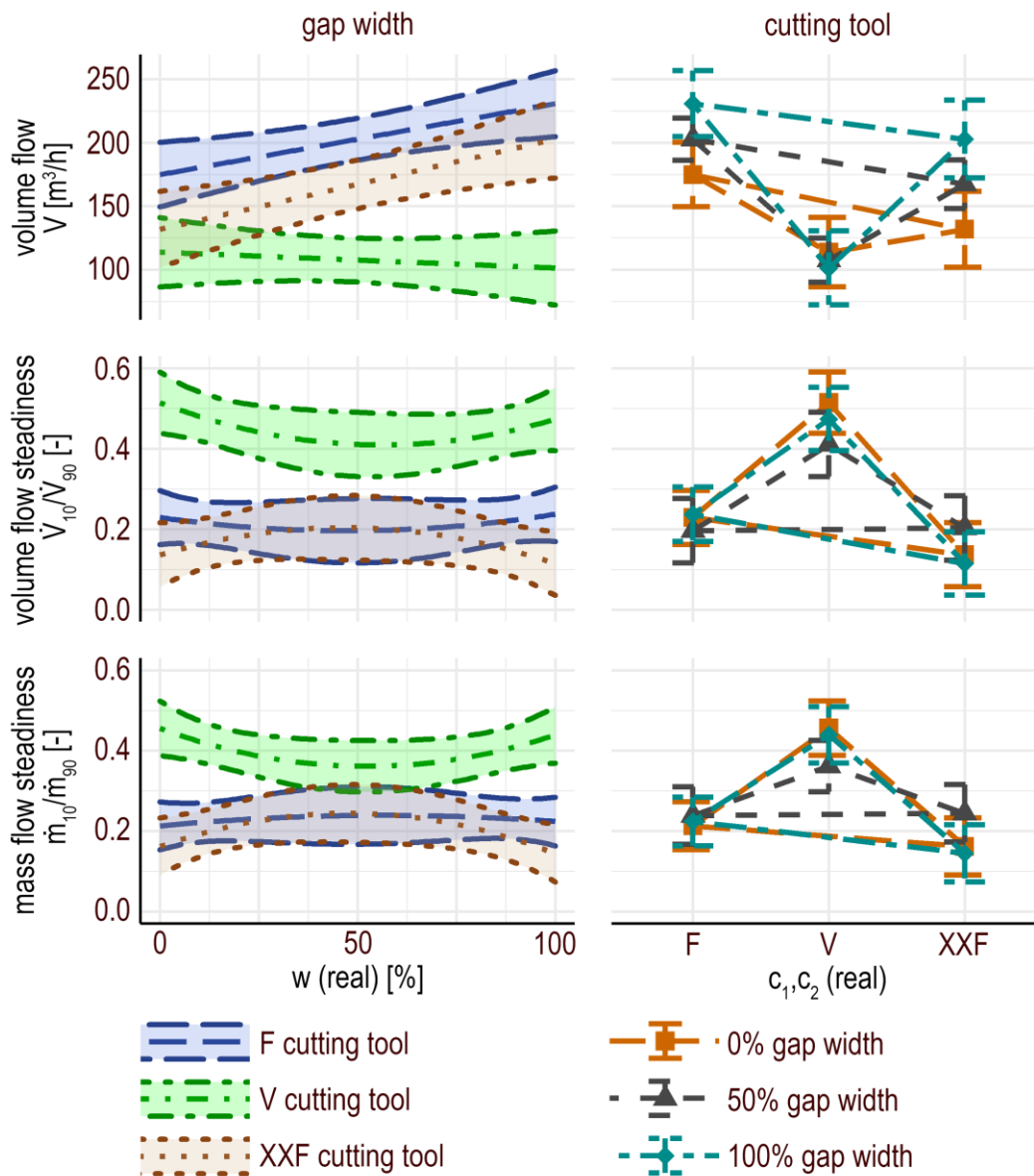
Figure 5: Interaction plots for the influence of the factors gap width (w) and cutting tool geometry (c) on the mean volume flow and the volume-related and mass-related throughput at different settings of the other factor (redesigned version of Fig. 4 in publication II – Khodier et al., 2021 – with absolute volume flows)
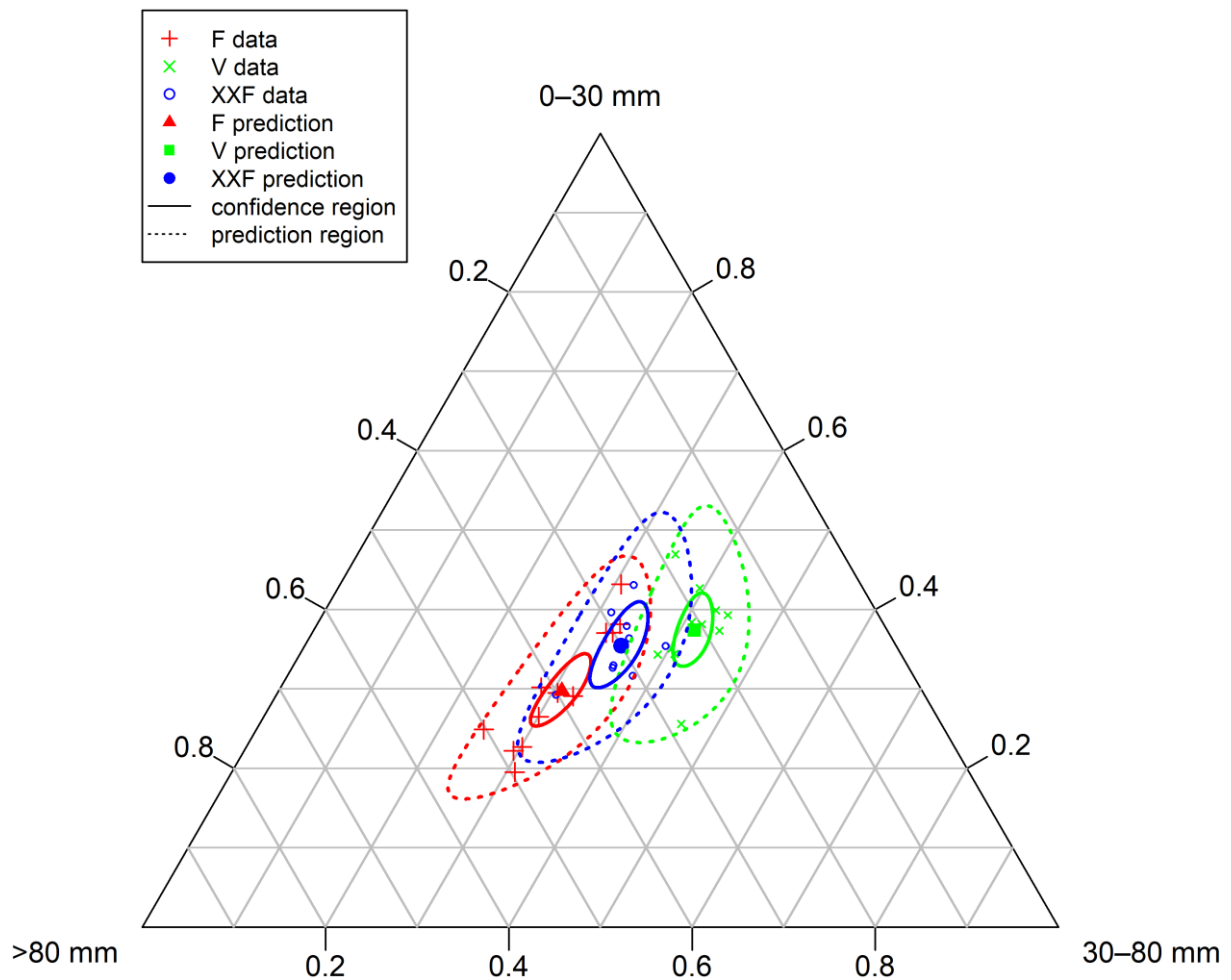
Figure 6: Prediction values and confidence and prediction regions for the particle size class distributions resulting from different cutting tool geometries (publication III: Khodier and Sarc, 2021)

4.5 Answer to research question 5

How can the particle size distribution of coarsely shredded mixed commercial waste, according to a screen, be determined in real-time, using state-of-the-art two-dimensional sensor data?

Partial least squares regression models, based on geometric descriptors for the particles, calculated from two-dimensional RGB images, turn out to be a promising approach for determining the particle size distribution of coarsely shredded mixed commercial waste, according to a screen. Separate models are needed for each material class, requiring material identification, e.g., through near-infrared sensors. Databases on typical area-specific particle masses can contribute the link from a distribution of particle counts to the usual presentation in terms of mass shares.

This research question is addressed in detail in **publication IV**.

Particle size distributions – in the context of waste processing – usually refer to the mass shares of particles of a specific size. Hence, two elements are necessary for their sensor-based determination: identifying a particle's size and its weight. For measurements of the individual particles, they must be present as a monolayer of single particles. An approach for obtaining the masses of the particles, in literature, consists of identifying a particle's material (e.g., using a near-infrared sensor) and multiplying average area-specific weights for this material with the particle's area, which is easily determined as a count of numbers of pixels. This thesis addresses the other component: calculating the particle size, as defined by a drum screen, from two-dimensional sensor data.

The applied approach is based on RGB images of the particles. It comprises three steps: First, the images are processed, using several optimizations and transformation algorithms, resulting in clear binary images, from which the contours of the particles can be identified. Next, a variety of geometric descriptors is calculated based on these outlines, including, for example, the minimal and maximal Feret diameters and the area of the smallest surrounding triangle. Considering these descriptors, and training data, containing the correct size class of the particles according to a drum screen, a partial least squares regression model is built. For characterizing new particles, their images are again processed, and the descriptors are calculated. Then, the model can be used to calculate the particle sizes from the descriptors.

In this thesis, the method was applied to particles of the fractions lightweight, and residual fraction, with the particle size classes 20–40 mm, 40–60 mm, 60–80 mm, >80 mm, and particles of the fractions wood, 3D-plastics, and paper and cardboard, of the same particle size classes, and additionally a class 10–20 mm.

The average share of correctly classified particles per size class by material ranges between 64.2% for paper and cardboard and 97.5% for the residual fraction. When determining the complete particle size distribution, some of the classification errors compensate each other – an effect that depends on the underlying present distribution. The resulting shares of the particle size classes were calculated, considering the size classes wrongly classified particles were assigned to, once assuming a uniform distribution, and once based on the particle size distribution of an actual mixed commercial waste, as analyzed in publication I. For the real distribution, the best results get as good as a maximum deviation of 1% for the shares of the size classes of wood, and the maximum average error per particle size class is 4% for the lightweight fraction.

5 CONCLUSION AND OUTLOOK

The thesis at hand aims to contribute to the methodological knowledge for performing and evaluating coarse-shredding experiments and mechanical waste processing experiments in general. Furthermore, it seeks to identify and quantify significant factor influences on these processes. It targets the optimized operation and real-time control of coarse shredders for mixed commercial waste, gaining insights into the effects of potential actuators. Moreover, it aims at presenting promising approaches for the sensor-based measurement of particle size distributions. The drawn conclusions and arisen future research questions are documented in the four following sections, divided into sampling, empirical modeling, the influence of shredding parameters, and sensor-based particle size measurement. Finally, a short presentation of the approved follow-up research program "Recycling and Recovery of Waste for Future" (ReWaste F), which will start in April 2021, is given.

5.1 Sampling

A procedure for sampling coarsely shredded mixed commercial waste was established, based on the theory of sampling and the Austrian standard ÖNORM S 2127. And for the first time, such a procedure was comprehensively evaluated theoretically and empirically – quantifying the general estimation error through a replication experiment. This combination provides novel insights on the nature of mixed commercial waste's heterogeneity, the corresponding sampling errors, and potential strategies for improving waste sampling.

Concluding from the theoretical evaluations, the defined procedure eliminates incorrect sampling errors almost completely. As for the correct sampling errors, the performed replication experiment indicates a high contribution from grouping and segregation errors, caused by a high level of distributional heterogeneity. Hence, dividing the target sample mass into more increments is likely to have a larger beneficial effect than raising the mass of the composite sample – an encouraging finding since an increased number of increments at a constant total mass does not affect the efforts for analyzing the composite sample.

Nevertheless, experimental validation of this conclusion is subject to future research. Finally, a database of typical weights of particles from specific particle size-material fractions was identified as a desirable target, providing a reference for the assumptions needed to predict the fundamental sampling error.

5.2 Empirical modeling

The behavior of a coarse shredder for mixed commercial waste was successfully modeled, applying univariate and multivariate multiple linear regression on data from a 32 runs Design of Experiments-based investigation, and describing the particle size distribution as isometric

log-ratio-transformed shares of a choice of particle size classes. The experiment is the first industry-scale parameter study on mechanical waste processing that uncompromisingly incorporates statistical principles on experimentation and analysis from the initial planning of the study to the final interpretation of the results. It seeks to function as a guide and benchmark for the scientific execution of mechanical waste processing experiments – demonstrating the extent of data distortion caused by the heterogeneity of the mixed waste while also proving the capability of the applied methods for nevertheless drawing reliable conclusions.

The applied approach has proven to be suitable for identifying and quantifying significant factor influences and calculating corresponding confidence and prediction regions, despite the difficult conditions introduced by the required industrial scale of the experiments, the heterogeneity of the waste, and the residual variance caused by sampling.

Concerning the modeling of particle size distributions, a distribution-independent, mathematically profound descriptive approach was found, suitably summarizing the data for regression modeling while preserving exactly the necessary level of information: for modeling purposes, the particle size distribution was described as isometric log-ratio-transformed shares of particle size classes. Limitations of the use of isometric log-ratios, caused by them introducing a relative scale, and not being defined, in case of the appearance of zero values, were found not to affect the specific case investigated in this thesis. Evaluating whether this is also the case for other particle size data from waste processing experiments and, if not, correctly applying existing approaches for handling these limitations, or developing new ones where necessary, are future research topics.

5.3   Influence of shredding parameters

The radial gap width, the shaft rotation speed, and the cutting tool geometry of a coarse shredder for mixed commercial waste were examined on the significance and magnitude of their influence on the average throughput and the throughput steadiness in terms of mass and volume, and on the specific energy demand, and the particle size distribution of the comminuted waste. The latter was summarized as the particle size classes >80 mm, 30–80 mm, and 0–30 mm.

The performed experimental study is the first of its kind, evaluating the influences of the three factors from 32 fully randomized runs of one hour, shredding and sampling around 700 metric tons of mixed commercial waste, while recording time-resolved mass flow and volume flow data of the comminution product.

All factors were identified as significant concerning the throughput behavior and energy demand. But only the cutting tool geometry significantly affected the shares of the defined particle size classes, which are relevant for producing solid recovered fuels. Consequently, the

results indicate optimal factor settings, which deviate from common practice: a larger gap width, for example, hence increases the throughput while decreasing the specific energy demand (which is both favorable), without negatively influencing the product quality.

The findings on the particle size distribution may, of course, differ for other choices of particle size classes. So further investigations may be needed, depending on the relevant particle size classes for a specific process. Moreover, the effects of the examined factors on material-specific particle size distributions were not investigated yet. They may provide more differentiated insights that are potentially relevant for the parametrization of shredders.

The observed effects are based on average values of an hour of operation. Concerning the self-adjustment of a smart processing plant to the waste, making the gap width and shaft rotation speed available for digital control appears promising. As for beneficially influencing the short-term behavior of shredders, for example, for smoothening their throughput, future research must show whether a dynamic adjustment of the same factors is a successful approach.

5.4 Sensor-based particle size measurement

Partial least squares regression models, based on geometric descriptors derived from two-dimensional images of single particles, were examined as an approach for the online measurement of the particle size distribution of shredded mixed commercial waste, according to drum screening.

The evaluation contributes the first work on the real-time measurement of mixed solid wastes' particle size distributions, defined in a way that reflects the particles' real process behavior. To the best of the doctoral candidate's knowledge, combining a variety of geometric descriptors of the projected surface for doing so is also a novelty.

The performance of the chosen method varies for different material classes. While it is quite promising for a residual fraction (classification accuracy of about 98%), the results for paper and cardboard are far from satisfactory (about 67% of the particles were correctly classified).

The suggested approach has proven to be promising. Still, several issues need to be addressed in future research: While classification errors partially cancel out, when measuring the total particle size distribution, this effect is distribution-dependent. Hence, enhancing the classification accuracy is necessary, particularly for those materials where the applied method performed poorly, potentially by replacing the regression with more sophisticated models, e.g., from machine learning.

Furthermore, not all kinds of materials were examined. For metals, for example, this was an issue of the contrast and quality of the images: the shapes of the particles could not be sufficiently identified. Therefore, further investigations on the determination of the shape are nec-

essary. These concern the fields of sensor technology and image processing, as well as considerations regarding the color and material of used conveyor belts, which constitute the background of the images.

Moreover, in this thesis, only particles with a minimum size of 10 mm (for wood, paper and cardboard, and 3D-plastics) or 20 mm (for the lightweight fraction and the residual fraction) were examined. For finer particles, the two following challenges must be addressed in future research activities: the sensor-based detection of the particles as objects, and the real-time computational handling of the tremendous number of single particles in such fine fractions.

Finally, investigations on the practicability of the approach for real-time measurements must also address the question of needed computational resources in general, considering the high number of particles per time and the number of subsequent steps of image processing and calculations. An approach for reducing the computational effort is digital "sampling", only evaluating a particular share of the detected particles.

5.5 Follow-up research program ReWaste F

The success of the ReWaste4.0 research program led to the application for and approval of funding for a follow-up research program entitled "Recycling and Recovery of Waste for Future." The consortium consists of four research institutions and fourteen industry partners, of which seven were also involved in ReWaste4.0. The approval of the four-year program, with a total budget of about 4.9 million Euros, ensures the continuation of the investigations, targeting the digitalized, smart, more effective and more efficient processing of solid waste, aiming to contribute to the transition towards a circular economy.

REFERENCES

Austrian Standards Institute, 2011. ÖNORM S 2127 Grundlegende Charakterisierung von Abfallhaufen oder von festen Abfällen aus Behältnissen und Transportfahrzeugen. Austrian Standards Institute, Vienna 13.030.01.

Berger, R., 2016. Die Digitalisierung in der GreenTech-Branche. Handlungsbedarfe für Unternehmen der Umwelttechnik und Ressourceneffizienz in Deutschland, Im Auftrag des: Bundesministeriums für Umwelt, Naturschutz, Bau und Reaktorsicherheit [Digitalisation in the GreenTech industry. Need for action for companies in environmental technology and resource efficiency in Germany, On behalf of: Federal Ministry of the Environment, Nature Conservation, Construction and Reactor Security].

BMK, 2020. Verordnung der Bundesministerin für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie über ein Abfallverzeichnis (Abfallverzeichnisverordnung 2020) [Ordinance of the Federal Minister for Climate Protection, Environment, Energy, Mobility, Innovation and Technology on a Waste Catalogue (Waste Catalogue Ordinance 2020)]: Abfallverzeichnisverordnung 2020 [Ordinance on a waste catalogue 2020].

Curtis, A., Küppers, B., Möllnitz, S., Khodier, K., Sarc, R., 2021. Digital material flow monitoring in waste processing - the relevance of material and throughput fluctuations. Waste Management 120, 687–697. https://doi.org/10.1016/j.wasman.2020.10.037.

European Union, 2008. Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain Directives.

European Union, 2018. Directive (EU) 2018/851 of the European parliament and of the council of 30 May 2018 amending Directive 2008/98/EC on waste.

Eurostat, 2020. Waste statistics. https://ec.europa.eu/eurostat/statistics-explained/index.php/Waste_statistics (accessed February, 2nd, 2021).

Feil, A., Pretz, T., 2018. Ungenutzte Potenziale in der Abfallaufbereitung [Unexploited potentials in waste processing], in: Pomberger, R. (Ed.), Recy & DepoTech 2018: Vorträge-Konferenzband zur 14. Recy & DepoTech-Konferenz. AVAW Eigenverlag.

German Federal Government, 2017. Verordnung über die Bewirtschaftung von gewerblichen Siedlungsabfällen und von bestimmten Bau- und Abbruchabfällen [Regulation on the management of commercial municipal waste and of certain construction and demolition waste].

Gundupalli, S.P., Hait, S., Thakur, A., 2017. A review on automated sorting of source-separated municipal solid waste for recycling. Waste Management 60, 56–74. https://doi.org/10.1016/j.wasman.2016.09.015.

Khodier, K., Curtis, A., Sarc, R., Lehner, M., O'Leary, P., Pomberger, R., 2019. Smart solid waste processing plant: vision and pathway, in: International Solid Waste Association (Ed.), Proceedings of ISWA word congress 2019, 7-9 October, Bilbao, Spain, pp. 194–195.

Khodier, K., Feyerer, C., Möllnitz, S., Curtis, A., Sarc, R., 2021. Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: coarse-shredding of commercial waste. Waste Management 121, 164–174. https://doi.org/10.1016/j.wasman.2020.12.015.

Khodier, K., Sarc, R., 2021. Distribution-independent empirical modeling of particle size distributions – coarse-shredding of mixed commercial waste. Processes 9 (3). https://doi.org/10.3390/pr9030414.

Khodier, K., Viczek, S.A., Curtis, A., Aldrian, A., O'Leary, P., Lehner, M., Sarc, R., 2020. Sampling and analysis of coarsely shredded mixed commercial waste. Part I: procedure, particle size and sorting analysis. Int. J. Environ. Sci. Technol. 17 (2), 959–972. https://doi.org/10.1007/s13762-019-02526-w.

Lee, H., Kwon, J.H., Kim, K.H., Cho, H.C., 2008. Application of DEM model to breakage and liberation behaviour of recycled aggregates from impact-breakage of concrete waste. Minerals Engineering 21 (11), 761–765. https://doi.org/10.1016/j.mineng.2008.06.007.

Luo, S., Xiao, B., Xiao, L., 2010. A novel shredder for municipal solid waste (MSW): influence of feed moisture on breakage performance. Bioresource technology 101 (15), 6256–6258. https://doi.org/10.1016/j.biortech.2010.02.067.

Luo, S.Y., Zhou, Y.M., Yi, C.J., 2011. A Novel Shredder for Municipal Solid Waste (*MSW*): Effect of Operation Mode and Feed Composition. AMR 308-310, 636–639. https://doi.org/10.4028/www.scientific.net/AMR.308-310.636.

Möllnitz, S., Küppers, B., Curtis, A., Khodier, K., Sarc, R., 2021. Influence of pre-screening on down-stream processing for the production of plastic enriched fractions for recycling from mixed commercial and municipal waste. Waste Management 119. https://doi.org/10.1016/j.wasman.2020.10.007.

Müller, W., Bockreis, A., 2015. Mechanical-Biological Waste Treatment and Utilization of Solid Recovered Fuels - State of the Art, in: Thomé-Kozmiensky, K.J., Thiel, S. (Eds.), Waste Management, vol. 5. TK Verlag Karl Thomé-Kozmiensky, Neuruppin, pp. 321–338.

Pawlowsky-Glahn, V., Egozcue, J.J., Tolosana-Delgado, R., 2015. Modeling and analysis of compositional data. John Wiley & Sons Inc, Chichester, West Sussex.

Pieper, C., Maier, G., Pfaff, F., Kruggel-Emden, H., Wirtz, S., Gruna, R., Noack, B., Scherer, V., Längle, T., Beyerer, J., Hanebeck, U.D., 2016. Numerical modeling of an automated optical belt sorter using the Discrete Element Method. Powder Technology 301, 805–814. https://doi.org/10.1016/j.powtec.2016.07.018.

Pomberger, R., 2008. Entwicklung von Ersatzbrennstoff für das HOTDISC-Verfahren und Analyse der abfallwirtschaftlichen Relevanz [Development of refuse-derived fuel for the HOTDISC process and analysis of the relevance for waste management]. Dissertation at Montanuniversitaet Leoben.

Sarc, R., Curtis, A., Kandlbauer, L., Khodier, K., Lorber, K.E., Pomberger, R., 2019. Digitalisation and intelligent robotics in value chain of circular economy oriented waste management - A review. Waste Management 95, 476–492. https://doi.org/10.1016/j.wasman.2019.06.035.

Schrör, H., 2011. Generation and treatment of waste in Europe 2008. Statistics in focus 44. https://ec.europa.eu/eurostat/web/products-statistics-in-focus/-/KS-SF-11-044.

Weißenbach, T.C., Hams, S., Walter, G., 2019. Grundlagen zum Thema Gewerbeabfälle in Bezug auf Echtzeitanalytik des Inputs [Fundamentals regarding commercial waste, concerning real-time input analytics]: Internal report within the research programme ReWaste4.0 at Montanuniversitaet Leoben. Leoben, Münster.

Zhang, Y., Kusch-Brandt, S., Gu, S., Heaven, S., 2019. Particle Size Distribution in Municipal Solid Waste Pre-Treated for Bioprocessing. Resources 8 (4), 166. https://doi.org/10.3390/resources8040166.

APPENDIX

A. Supplementary material of Publication II

Efficient derivation of significant results from mechanical processing experiments with mixed solid waste: coarse-shredding of commercial waste

K. Khodier[a], C. Feyerer[c], S. Möllnitz[b], A. Curtis[b], R. Sarc[b]*

[a]*Chair of Industrial Environmental Protection and Process Engineering,*

Montanuniversitaet Leoben, Leoben, Austria

[b]*Chair of Waste Processing Technology and Waste Management,*

Montanuniversitaet Leoben, Leoben, Austria

[c]*Komptech GmbH, Frohnleiten, Austria*

*Corresponding author.

Supplementary material

Figure S1: Exemplary photo of the used waste
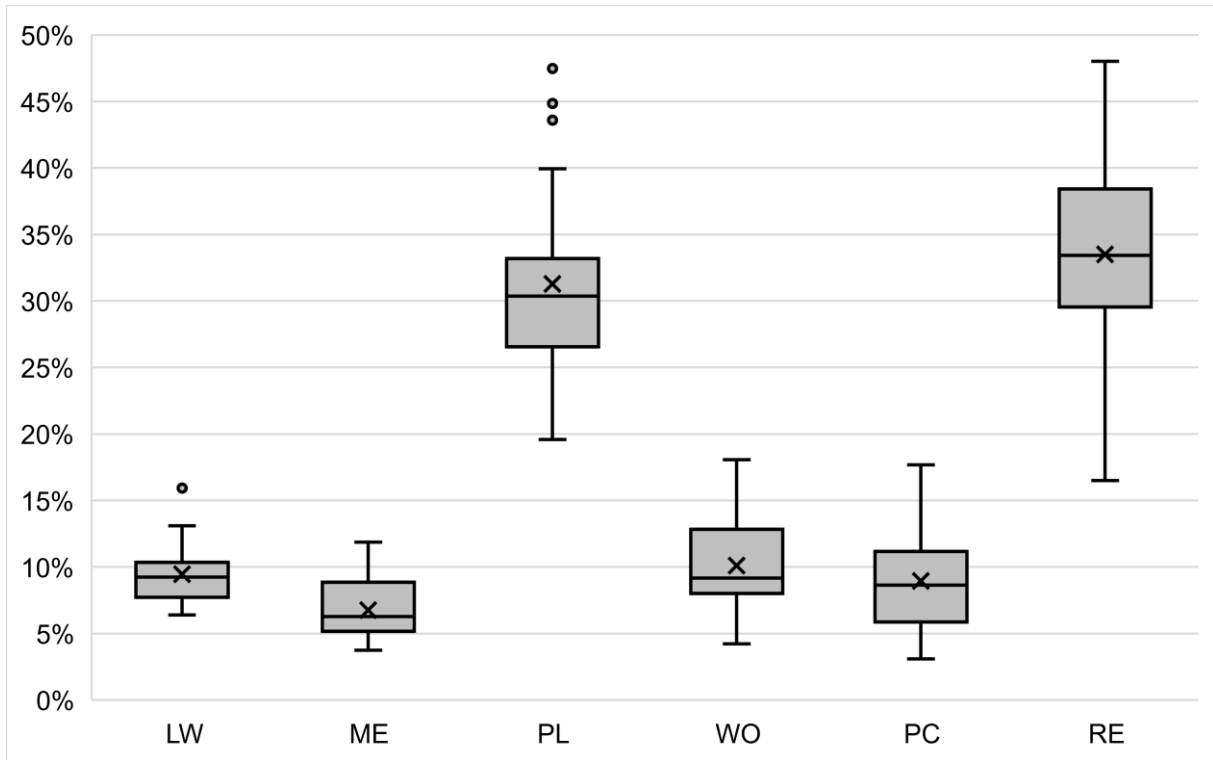
Figure S2: Box and whisker plot, showing the results of the sorting analyses of samples from runs 1-31. LW is a lightweight fraction, obtained through wind-sifting of all material >10 mm and mainly consists of plastic films and foams. The heavyweight fraction was then sorted into ME (metal), PL (plastic), WO (wood), PC (paper and cardboard), RE (residual fration). Material <10 mm was also assighned to RE. The analyses were performed without prior drying.

Table S1: Experimental design (non-coded) and results

	Design				Results				
Run actual	Run planned	$w[\%]$	$s[\%]$	$c[-]$	$\dot{V}[-]$	$\dot{m}[-]$	$\dot{V}_{10}/\dot{V}_{90}[-]$	$\dot{m}_{10}/\dot{m}_{90}[-]$	$E[\text{Lt}^{-1}]$
1	1	0	60	XXF	0.61	0.71	0.21	0.16	2.5
2	2	100	100	XXF	1.19	1.27	0.07	0.14	1.9
3	3	0	60	V	0.49	0.57	0.45	0.45	4.4
4	4	50	80	V	0.41	0.34	0.31	0.29	8.9
5	5	0	100	XXF	0.92	0.92	0.06	0.08	2.6
6	6	100	60	V	0.54	0.55	0.44	0.39	4.2
7	7	50	80	F	0.99	0.89	0.17	0.21	2.3
8	8	100	60	F	1.14	0.91	0.19	0.19	1.8
9	9	0	100	F	1.15	1.03	0.18	0.14	2.1
10	10	100	100	F	1.63	1.15	0.18	0.16	1.8
11	11	100	60	XXF	1.02	1.18	0.15	0.14	1.4
12	12	20	70	V	0.67	0.60	0.44	0.42	4.2
13	13	0	100	XXF	0.94	1.15	0.10	0.21	1.9
14	14	0	60	F	0.92	0.65	0.25	0.28	2.8
15	15	20	70	F	1.00	1.15	0.24	0.27	1.7
16	17	50	60	XXF	0.93	0.91	0.26	0.32	2.0
17	20	40	80	XXF	1.19	1.32	0.24	0.27	1.9
18	21	70	70	F	1.21	1.02	0.23	0.26	1.8
19	22	0	80	F	1.48	1.48	0.31	0.28	1.5
20	23	50	60	F	1.29	0.82	0.24	0.29	2.1
21	24	30	90	V	0.51	0.76	0.09	0.33	4.4
22	16	0	80	V	0.96	1.18	0.58	0.45	2.3
23	18	100	100	V	0.79	1.02	0.41	0.40	3.2
24	19	100	100	V	0.78	0.88	0.49	0.47	4.0
25	25	100	60	F	1.35	1.26	0.35	0.35	1.4
26	26	50	100	XXF	1.23	1.33	0.11	0.14	1.9
27	27	50	60	V	0.57	0.67	0.46	0.43	3.5
28	28	70	70	V	0.76	1.11	0.57	0.46	2.3
29	29	0	100	V	0.89	0.93	0.54	0.48	3.5
30	30	100	80	XXF	1.49	1.66	0.13	0.15	1.4
31	31	100	100	F	1.81	1.42	0.23	0.20	1.6
32	32	0	100	F	1.13	1.18	0.13	0.11	2.0
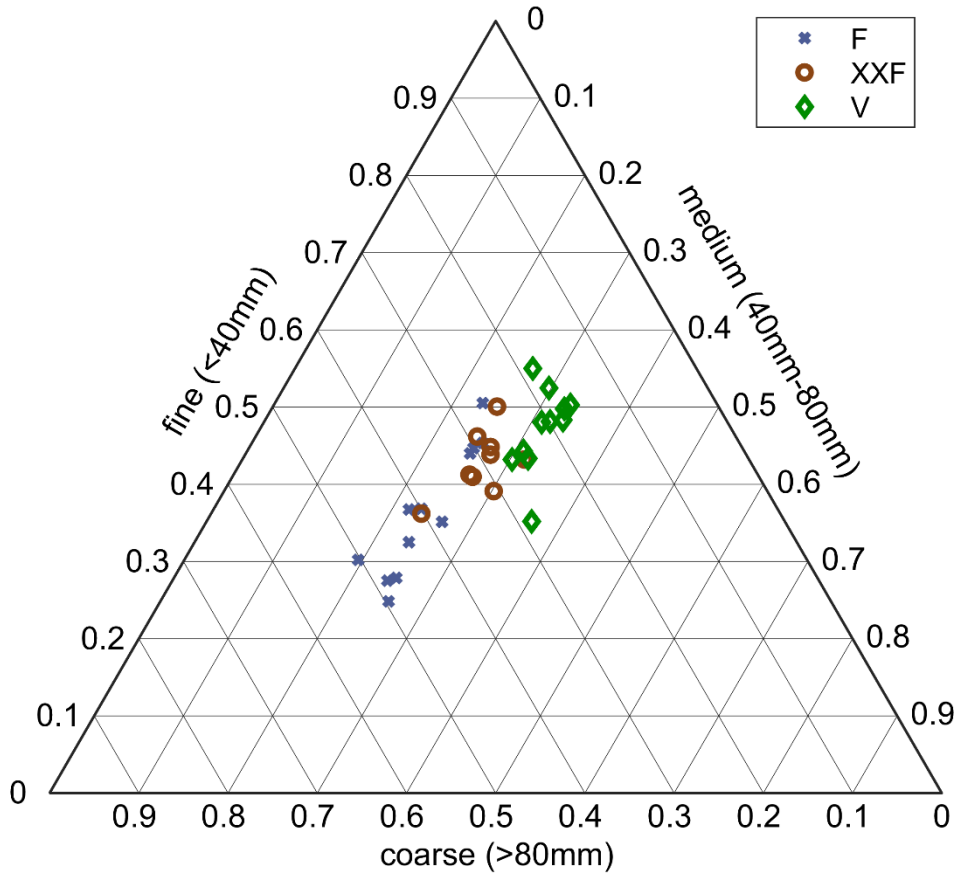
23

24   Figure S3: Particle size distributions (mass share) of the products of the three cutting tool geometries

B. Supplementary material of Publication III

Distribution-Independent Empirical Modeling of Particle Size Distributions – Coarse-Shredding of Mixed Commercial Waste

Supplementary Material: Model Finding and Analysis

K. Khodier[1], R. Sarc[2]*

[1]Department of Environmental and Energy Process Engineering; Chair of Process Technology and Industrial Environmental Protection, Montanuniversitaet Leoben - Franz-Josef-Straße 18, 8700 Leoben, Austria

[2]Department of Environmental and Energy Process Engineering; Chair of Waste Processing Technology and Waste Management, Montanuniversitaet Leoben - Franz-Josef-Straße 18, 8700 Leoben, Austria

*Corresponding author. Tel.: +43 3842 402-5105; E-mail address: renato.sarc@unileoben.ac.at

In this document, the code for the data analysis is provided, and the process of finding and analyzing the model of the multivariate multiple linear regression is documented.

Preparations

Preparing the Environment

The required packages are loaded:

- compositions: provides many tools for handling compositional data
- regr: is needed for calculating the significances of factors and interactions during model finding
- MPV: is used for calculating PRESS, to calculate $R^2_{predict}$
- MVN: is used for testing on multivariate normality

```
In [1]:  library(compositions)
```

```
Welcome to compositions, a package for compositional data analysis.
Find an intro with "? compositions"


Attaching package: 'compositions'


The following objects are masked from 'package:stats':

    cor, cov, dist, var


The following objects are masked from 'package:base':

    %*%, norm, scale, scale.default
```

```
In [2]:   library(regr)
```

Loading required package: plgraphics

Registered S3 methods overwritten by 'plgraphics':
  method   from
  nobs.survreg survival
  nobs.coxph   survival

Registered S3 methods overwritten by 'regr':
  method   from
  add1.default   stats
  add1.mlm   stats
  drop1.default stats
  drop1.mlm   stats
  drop1.multinom nnet

```
In [3]:   library(MPV)
```

Warning message:
"package 'MPV' was built under R version 4.0.3"
Loading required package: lattice

Loading required package: KernSmooth

KernSmooth 2.23 loaded
Copyright M. P. Wand 1997-2009

```
In [4]:   library(MVN)
```

Registered S3 methods overwritten by 'car':
  method   from
  influence.merMod   lme4
  cooks.distance.influence.merMod lme4
  dfbeta.influence.merMod   lme4
  dfbetas.influence.merMod   lme4

Registered S3 method overwritten by 'GGally':
  method from
  +.gg   ggplot2

sROC 0.1-2 loaded

The versions of R and the loaded packages are printed, for the purpose of documentation:

```
In [5]:   sessionInfo()
```

R version 4.0.2 (2020-06-22)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows 10 x64 (build 17134)

Matrix products: default

locale:
[1] LC_COLLATE=German_Austria.1252 LC_CTYPE=German_Austria.1252
[3] LC_MONETARY=German_Austria.1252 LC_NUMERIC=C
[5] LC_TIME=German_Austria.1252

attached base packages:
[1] stats   graphics grDevices utils   datasets methods   base

other attached packages:
[1] MVN_5.8   MPV_1.56   KernSmooth_2.23-17 lattice_0.20-41

```
 [5] regr_1.1              plgraphics_1.1      compositions_2.0-1

loaded via a namespace (and not attached):
  [1] minqa_1.2.4            colorspace_1.4-1      mvoutlier_2.0.9
  [4] modeltools_0.2-23     ellipsis_0.3.1        class_7.3-17
  [7] rio_0.5.16            mclust_5.4.6          IRdisplay_0.7.0
 [10] pls_2.7-3             base64enc_0.1-3       rstudioapi_0.11
 [13] cvTools_0.3.2         MatrixModels_0.4-1    flexmix_2.3-15
 [16] mvtnorm_1.1-1         ranger_0.12.1         xml2_1.3.2
 [19] sROC_0.1-2            splines_4.0.2         leaps_3.1
 [22] mnormt_2.0.2          robustbase_0.93-6     knitr_1.29
 [25] IRkernel_1.1.1        jsonlite_1.7.1        nloptr_1.2.2.2
 [28] robCompositions_2.2.1 kernlab_0.9-29       cluster_2.1.0
 [31] rrcov_1.5-5           compiler_4.0.2        httr_1.4.2
 [34] Matrix_1.2-18         htmltools_0.5.0       quantreg_5.61
 [37] tools_4.0.2           gtable_0.3.0          glue_1.4.2
 [40] dplyr_1.0.2           Rcpp_1.0.5            carData_3.0-4
 [43] cellranger_1.1.0      zCompositions_1.3.4  vctrs_0.3.4
 [46] sgeostat_1.0-27       nlme_3.1-148          conquer_1.0.2
 [49] fpc_2.2-7             psych_2.0.8           lmtest_0.9-38
 [52] tensorA_0.36.1        xfun_0.17             laeken_0.5.1
 [55] stringr_1.4.0         openxlsx_4.1.5        lme4_1.1-23
 [58] rvest_0.3.6           lifecycle_0.2.0       statmod_1.4.34
 [61] DEoptimR_1.0-8        MASS_7.3-51.6         zoo_1.8-8
 [64] scales_1.1.1          VIM_6.0.0             hms_0.5.3
 [67] parallel_4.0.2        SparseM_1.78          RColorBrewer_1.1-2
 [70] curl_4.3              NADA_1.6-1.1          ggplot2_3.3.2
 [73] reshape_0.8.8         stringi_1.4.6         nortest_1.0-4
 [76] pcaPP_1.9-73          e1071_1.7-3           energy_1.7-7
 [79] boot_1.3-25           zip_2.1.1             truncnorm_1.0-8
 [82] chron_2.3-56          repr_1.1.0            prabclus_2.3-2
 [85] rlang_0.4.7           pkgconfig_2.0.3       moments_0.14
 [88] matrixStats_0.56.0    evaluate_0.14         purrr_0.3.4
 [91] tidyselect_1.1.0      GGally_2.0.0          plyr_1.8.6
 [94] magrittr_1.5          R6_2.4.1              generics_0.0.2
 [97] pbdZMQ_0.3-3.1        pillar_1.4.6          haven_2.3.1
[100] foreign_0.8-80        survival_3.1-12       abind_1.4-5
[103] sp_1.4-2              nnet_7.3-14           tibble_3.0.3
[106] bayesm_3.1-4          crayon_1.3.4          car_3.0-9
[109] uuid_0.1-4            tmvnsim_1.0-2         rmarkdown_2.3
[112] grid_4.0.2            readxl_1.3.1          data.table_1.13.0
[115] forcats_0.5.0         diptest_0.75-7        vcd_1.4-7
[118] digest_0.6.25         webshot_0.5.2         tidyr_1.1.2
[121] stats4_4.0.2          munsell_0.5.0         viridisLite_0.3.0
[124] kableExtra_1.2.1
```

Loading the Data

The matrix of the independent variables $\mathbf{X}$ is loaded into a variable "X" from a .csv file. The file contains three columns, which contain the settings for the gap width (w), and the shaft rotation speed (s) in coded form. The third column contains the cutting tool geometry (c) in non-coded form. The rows are ordered according to the order the runs were performed.

The matrix of the dependent variable $\mathbf{Y}$ is loaded into a variable "Y" from a .csv file. The file also contains three columns. They contain the shares of the fractions >80 mm, 30-80 mm, and 0-30 mm, which are named "gr80", "y30-80", "y0-30".

In [6]:
```r
X<-read.table(file="X.csv", header=TRUE)
print(X)
```

```
     w    s   c
1  -1.0 -1.0 XXF
2   1.0  1.0 XXF
3  -1.0 -1.0   V
4   0.0  0.0   V
```

```
 5  -1.0   1.0 XXF
 6   1.0  -1.0   V
 7   0.0   0.0   F
 8   1.0  -1.0   F
 9  -1.0   1.0   F
10   1.0   1.0   F
11   1.0  -1.0 XXF
12  -0.6  -0.5   V
13  -1.0   1.0 XXF
14  -1.0  -1.0   F
15  -0.6  -0.5   F
16   0.0  -1.0 XXF
17  -0.2   0.0 XXF
18   0.4  -0.5   F
19  -1.0   0.0   F
20   0.0  -1.0   F
21  -0.4   0.5   V
22  -1.0   0.0   V
23   1.0   1.0   V
24   1.0   1.0   V
25   1.0  -1.0   F
26   0.0   1.0 XXF
27   0.0  -1.0   V
28   0.4  -0.5   V
29  -1.0   1.0   V
30   1.0   0.0 XXF
31   1.0   1.0   F
32  -1.0   1.0   F
```

In [7]:
```r
Y<-read.table(file="Y.csv",header=TRUE)
names(Y)<-c("gr80","y30-80","y0-30")
print(Y)
```

```
     gr80 y30-80  y0-30
1  0.3210 0.3490 0.3300
2  0.2898 0.3135 0.3968
3  0.1833 0.4431 0.3735
4  0.2840 0.4602 0.2558
5  0.2866 0.3496 0.3638
6  0.1646 0.4424 0.3930
7  0.3993 0.3057 0.2950
8  0.4137 0.2843 0.3020
9  0.2616 0.3067 0.4317
10 0.5029 0.2482 0.2489
11 0.2482 0.3208 0.4310
12 0.2469 0.4088 0.3442
13 0.4023 0.3048 0.2929
14 0.4837 0.2942 0.2221
15 0.2878 0.3306 0.3815
16 0.2517 0.3942 0.3541
17 0.2816 0.3392 0.3792
18 0.3844 0.3247 0.2908
19 0.3013 0.3277 0.3711
20 0.4958 0.3088 0.1954
21 0.1985 0.4200 0.3815
22 0.2086 0.4076 0.3839
23 0.2471 0.4021 0.3508
24 0.1781 0.3951 0.4268
25 0.4348 0.3002 0.2650
26 0.3067 0.3767 0.3166
27 0.2657 0.3908 0.3434
28 0.1834 0.3470 0.4697
29 0.1744 0.4266 0.3990
30 0.3232 0.3498 0.3270
31 0.4719 0.3012 0.2269
32 0.3086 0.3207 0.3707
```

Preparing the Data

For the purpose of simplicity, the data from X is extracted into individual variables. The gap width w and the shaft rotation speed s are saved as "w" and "s". Their squares are also saved as variables "w2" and "s2" for easier handling in modeling. The cutting tool geometry c is saved as a variable "c", defining, that it contains factor levels. Furthermore, it is defined, that sum contrasts are used for coding c. Finally, the contrast matrix of c is printed.

```
In [8]:  w<-X$w
         w2<-w^2
         s<-X$s
         s2<-s^2
         c<-factor(X$c)
         contrasts(c)<-"contr.sum"
         contrasts(c)
```

A matrix: 3 × 2

of type dbl

F	1	0
V	0	1
XXF	-1	-1

Next, it is defined, that Y contains compositional data. Then, the isometric log-ratio transformations of Y are saved as "Y.ilr", using the default ilr base. The names of the columns of Y.ilr are then named "ilr1" and "ilr2". Finally, the ilr base of Y.ilr is printed.

```
In [9]:  Y<-acomp(Y)
         Y.ilr<-ilr(Y)
         names(Y.ilr)<-c("ilr1","ilr2")
         ilrBase(Y)
```

A matrix: 3 × 2 of type dbl

1	-0.7071068	-0.4082483
2	0.7071068	-0.4082483
3	0.0000000	0.8164966

Data Analysis

Visualization

To get a first impression, dependencies of the compositional and ilr-transformed data on w, s, and c are visualized.

```
In [10]:  pairwisePlot(cbind(w,s,c),Y)
```

```
Warning message in .S3methods(generic.function, class, envir):
"generic function 'panel' dispatches methods for generic 'plot'"
```
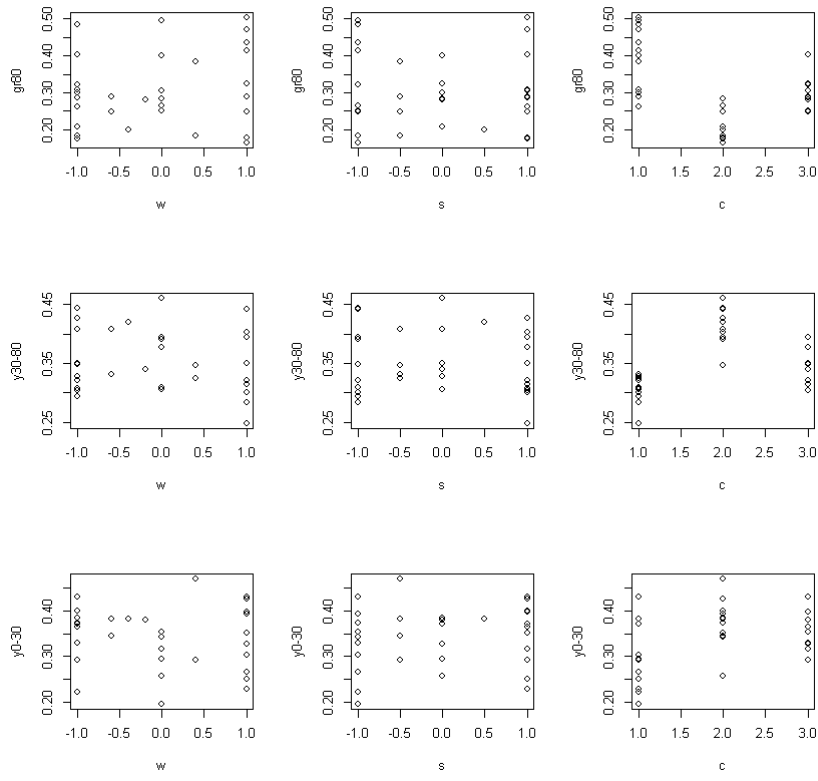
```
pairwisePlot(cbind(w,s,c),Y.ilr)
```

```
Warning message in .S3methods(generic.function, class, envir):
"generic function 'panel' dispatches methods for generic 'plot'"
```
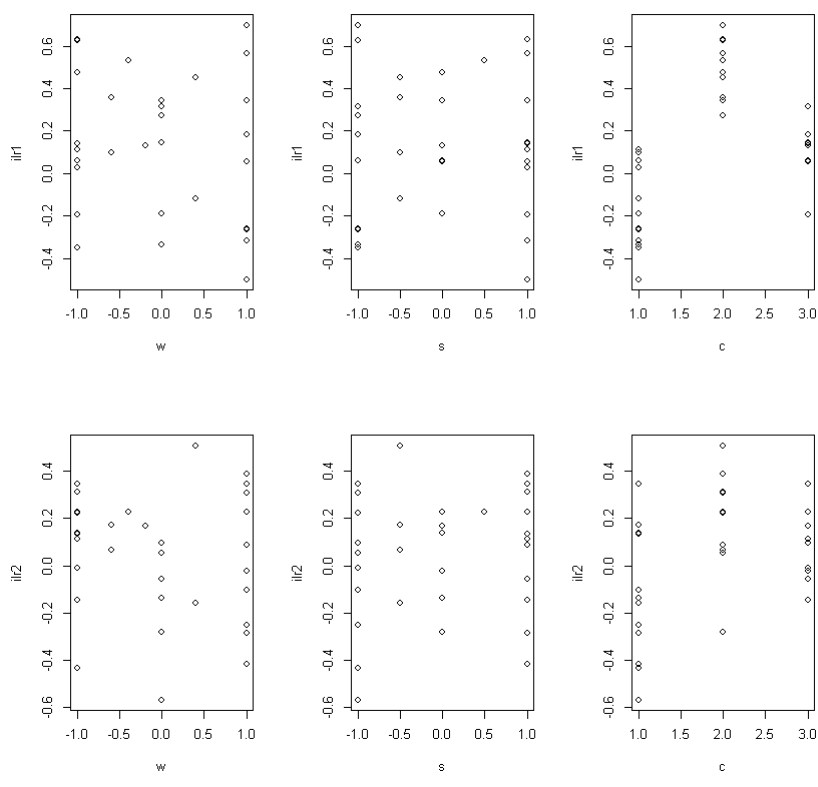

Based on the plot, dependencies on the cutting tool geometry are visible. Their significance is subject to further analyses. No immedeately visible dependencies on the gap width and shaft rotation speed can be seen.

Model Choice

A proper model is chosen, starting with the reduced cubic design model, according to Equation (14) in the paper:

$$\hat{y}^{(r)} = \sum_{j=0}^{2} \sum_{k=0}^{2-j} \sum_{m=0}^{1} \sum_{n=0}^{1-m} \left(K_{jkmn}^{(r)} w^j s^k c_1^m c_2^n \right)$$

The model is saved in a variable "model".

In [12]:
```
model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+s2+w:s+w:c+s:c+w2:c+s2:c+w:s:c,data=Y.ilr)
```

In the following, the significance of all factors and interactions, which can be dismissed without violating the model's hierarchy is evaluated. This is done using the function "drop1()" from the "regr" package, which calculates each dropable factors' p-value, for the case where all other factors are kept. Since "drop1()" does not identify s2 and w2 as higher order terms of s and w, the hierarchy must additionally be checked manually. As long as dropable factors or interactions with a p-value higher than p=0.1 are present, the term with the highest p-value is droped, and the p-values are re-calculated.

In [13]:
```
drop1(model,test="Pillai")
```

A data.frame: 4 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
c:w2	0.3565918	1.5188817	4	28	0.2236688
c:s2	0.2486506	0.9938361	4	28	0.4272063
w:s:c	0.1884995	0.7283999	4	28	0.5801124
<total>	1.4072946	2.3743574	28	28	0.0127443

All printed terms are dropable, considering hierarchy. Hence "w:s:c" is dropped, as it has the highest p-value.

In [14]:
```
model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+s2+w:s+w:c+s:c+w2:c+s2:c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 6 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w:s	0.3103325	3.3748064	2	15	0.06163082
w:c	0.3102038	1.4685975	4	32	0.23477507
s:c	0.1634753	0.7121070	4	32	0.58974894
c:w2	0.3436120	1.6595724	4	32	0.18366090
c:s2	0.2045822	0.9115748	4	32	0.46905331
<total>	1.3549123	2.1003536	32	32	0.01971546

"w:c" and "s:c" must be kept for the model's hierarchy. "c:s2" is the remaining term with the highest p-value and is therfore dropped.

In [15]:
```
model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+s2+w:s+w:c+s:c+w2:c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 6 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
s2	0.02908321	0.2546122	2	17	0.77812093
w:s	0.25141769	2.8547968	2	17	0.08531713
w:c	0.31387435	1.6753610	4	36	0.17704982
s:c	0.14499921	0.7034999	4	36	0.59473494
c:w2	0.35409505	1.9362330	4	36	0.12549707
<total>	1.29160622	1.8232885	36	36	0.03783630

"w:c" must be kept for the model's hierarchy. "s2" is the remaining term with the highest p-value and is therefore dropped.

```
In [16]:   model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+w:s+w:c+s:c+w2:c,data=Y.ilr)
           drop1(model,test="Pillai")
```

A data.frame: 5 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w:s	0.2445743	2.9138119	2	18	0.08011723
w:c	0.3285391	1.8673012	4	38	0.13623036
s:c	0.1422602	0.7274817	4	38	0.57873901
c:w2	0.3504631	2.0183844	4	38	0.11136184
<total>	1.2763102	1.7636150	38	38	0.04218230

"w:c" must be kept for the model's hierarchy. "s:c" is the remaining term with the highest p-value and is therefore dropped.

```
In [17]:   model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+w:s+w:c+w2:c,data=Y.ilr)
           drop1(model,test="Pillai")
```

A data.frame: 4 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w:s	0.1935762	2.400428	2	20	0.11631433
w:c	0.3285402	2.063868	4	42	0.10273368
c:w2	0.3416119	2.162898	4	42	0.08980521
<total>	1.2264776	1.585575	42	42	0.06966111

"w:c" must be kept for the model's hierarchy. "w:s" is the remaining term with the highest p-value and is therefore dropped.

```
In [18]:   model<-lm(cbind(ilr1,ilr2)~w+s+c+w2+w:c+w2:c,data=Y.ilr)
           drop1(model,test="Pillai")
```

A data.frame: 4 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
s	0.0488256	0.5389851	2	21	0.5911962
w:c	0.2731460	1.7399305	4	44	0.1583155
c:w2	0.2965543	1.9149992	4	44	0.1247918
<total>	1.1477190	1.3466439	44	44	0.1636184

"w:c" must be kept for the model's hierarchy. "s" is the remaining term with the highest p-value and is therefore dropped.

In [19]:
```
model<-lm(cbind(ilr1,ilr2)~w+c+w2+w:c+w2:c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 3 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w:c	0.2734450	1.821325	4	46	0.1409107
c:w2	0.2834389	1.898883	4	46	0.1267175
<total>	1.1337923	1.308915	46	46	0.1823114

"w:c" must be kept for the model's hierarchy. "c:w2" is the remaining term with the highest p-value and is therefore dropped.

In [20]:
```
model<-lm(cbind(ilr1,ilr2)~w+c+w2+w:c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 3 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w2	0.1009006	1.346689	2	24	0.2790568
w:c	0.2374396	1.683911	4	50	0.1684308
<total>	1.0880683	1.193147	50	50	0.2673526

Both terms are dropable, considering model hierarchy. "w2" is dropped, being the term with the highest p-value.

In [21]:
```
model<-lm(cbind(ilr1,ilr2)~w+c+w:c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 2 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w:c	0.235857	1.738034	4	52	0.1556973
<total>	1.009897	1.019993	52	52	0.4716869

"w:c" is the only dropable term. Since its p-value is higher than 0.1 it is dropped.

In [22]:
```
model<-lm(cbind(ilr1,ilr2)~w+c,data=Y.ilr)
```

```
drop1(model,test="Pillai")
```

A data.frame: 3 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
w	0.0948458	1.4145858	2	27	2.604682e-01
c	0.8619933	10.6044255	4	56	1.817275e-06
<total>	0.8666652	0.7647036	56	56	8.408135e-01

"w" is the term with the higher p-value. Since its higher than 0.1, it is dropped.

In [23]:
```
model<-lm(cbind(ilr1,ilr2)~c,data=Y.ilr)
drop1(model,test="Pillai")
```

A data.frame: 2 × 5

	Pillai	F.stat	dfnum	dfden	p.value
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
c	0.8479475	10.672464	4	58	1.501489e-06
<total>	0.8479475	0.736032	58	58	8.769300e-01

The remaining term "c" is highly significant, with a p-value much smaller than 0.1. Hence, the final model is found:

$$\hat{y}^{(r)} = K_{0000} + K_{0010}c_1 + K_{0001}c_2$$

Model Analysis

First, three different coefficients of determinations are calculated. The regular coefficient of determination R^2 ("R2"), the adjusted coefficient of determination R^2_{adj} ("R2.adj"), and the prediction coefficient of determination R^2_{pred} ("R2.pred"), which is calculated, based on Allen's $PRESS$, using the "PRESS" function from the "MPV" package, and the total sum of squares ("SST").

In [24]:
```
(R2<-R2(model, adjust=FALSE))
```

0.571286609899341

In [25]:
```
(R2.adj<-R2(model, adjust=TRUE))
```

0.541720169202744

In [26]:
```
SST<-sum((Y.ilr-mean(Y.ilr))^2)
PRESS<-PRESS(model)
(R2.pred<-(1-(PRESS/SST)))
```

0.482605122103295

The model covers 57% of the observed variance, with an adjusted coefficient of determination of 0.54. While these values do not seem very high at a first glance, they are not bad, considering the expectable residual error due to changes in the waste. The prediction coefficient of determination is close to the adjusted coefficient of determination. Consequently, overfitting is

not very likely. This makes sense, as the number of data is considerably higher than the number of model coefficients.

Next, the residuals are analyzed. Fist the residuals of each ilr-coordinate are saved as "ilr1.resid" and "ilr2.resid". Then, they are grahpically checked for their univariate normality, using quantile-quantile-plots.

In [27]:
```
ilr1.resid<-resid(model)[,1]
ilr2.resid<-resid(model)[,2]
```
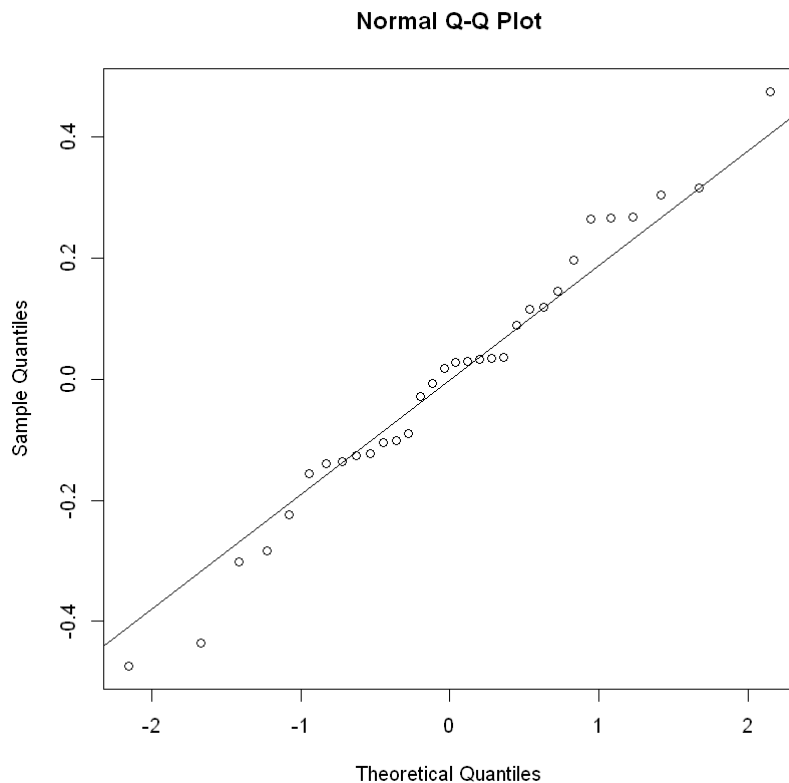
The plot for ilr1:

In [28]:
```
qqnorm(ilr1.resid)
qqline(ilr1.resid)
```
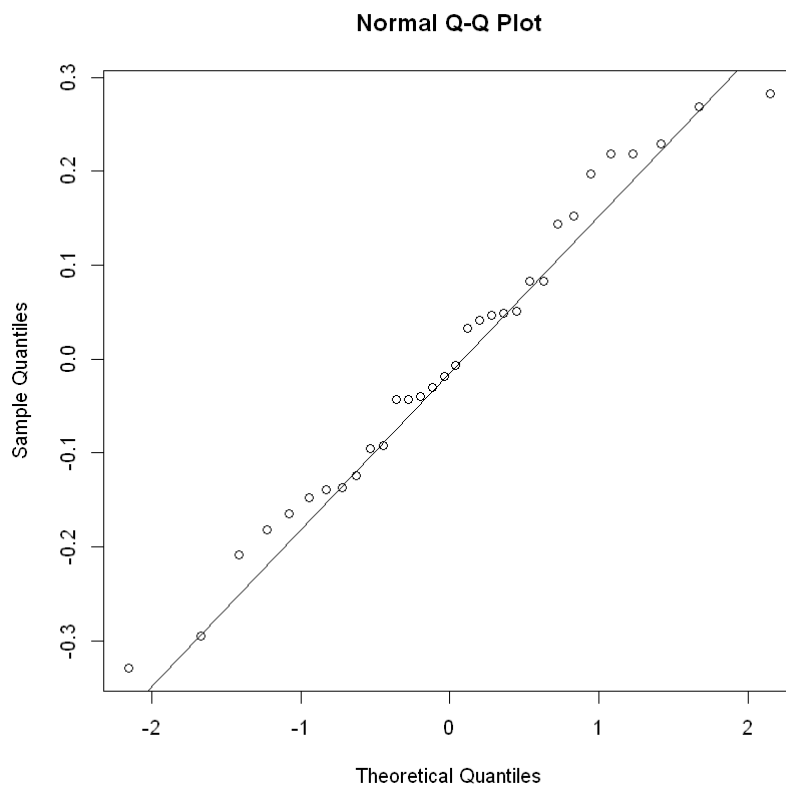

The plot does not indicate deviations from normality, which is good. The plot for ilr2:

In [29]:
```
qqnorm(ilr2.resid)
qqline(ilr2.resid)
```

Normal Q-Q Plot

The plot also does not show deviatios from normality. Finally, the residuals are tested on multivariate normaility, based ond Mardia's Skewness and Mardia's Kurtosis. The used function "mvn" from the MVN package also tests the residuals on univariate normality, based on the Shapiro-Wilk test.

```
In [30]:   mvn(resid(model))
```

$multivariateNormality

A data.frame: 3 × 4

Test	Statistic	p value	Result
<chr>	<fct>	<fct>	<chr>
Mardia Skewness	1.1144290431284	0.891976172305813	YES
Mardia Kurtosis	-0.3772196099597	0.706010408281269	YES
MVN	NA	NA	YES

$univariateNormality

A data.frame: 2 × 5

	Test	Variable	Statistic	p value	Normality
	<I<chr>>	<I<chr>>	<I<chr>>	<I<chr>>	<I<chr>>
1	Shapiro-Wilk	ilr1	0.9759	0.6739	YES
2	Shapiro-Wilk	ilr2	0.9834	0.8905	YES

$Descriptives

A data.frame: 2 × 10

	n	Mean	Std.Dev	Median	Min	Max	
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<
ilr1	32	4.445229e-18	0.1610418	-0.01293467	-0.3292027	0.2825931	-0.127
ilr2	32	-5.421011e-19	0.2193393	0.02178508	-0.4741227	0.4746113	-0.128

The tests confirm univariate and multivariate normality. Hence, the tests from the MANOVA are valid and consequently the model is accepted.

Model Representation

In the following, the confidence and prediction bands for the model will be visualized, based on the code described by van den Boogaart and Tolosana-Delgado [40]. First, a function "getModelMatrix" is defined, which outputs the X matrix of the model in a representation that consideres applied contrasts.

In [31]:
```
getModelMatrix<-function(object,newdata=NULL,na.action=na.pass) {
  if(is.null(newdata))
    return(model.matrix(object))
  Terms<-delete.response(terms(object))
  mf<-model.frame(Terms,newdata,na.action=na.action,xlev=object$xlevels)
  if(!is.null(cl<-attr(Terms,"dataClasses")))
    .checkMFClasses(cl,mf)
  model.matrix(Terms,mf,contrasts.arg=object$contrasts)
}
```

Next, the prediction value and the confidence and prediction ellipses for the F cutting tool are calculated and represented in the back-transformed Euclidian space. For doing so, first, new data is defined:

In [32]:
```
F.newdata<-data.frame(c="F")
```

Then, the prediction values are calculated, using the "predict" function, and backtransforming the results using the "ilrInv()". The results are saved as "F.prediction".

In [33]:
```
(F.prediction<-ilrInv(predict(model,newdata=F.newdata),orig=Y))
```

```
      gr80      y30-80      y0-30
0.3934214 0.3092926 0.2972859
attr(,"class")
[1] acomp
```

Afterwards, the residual variance "varEpsilon" is calculated:

In [34]:
```
(varEpsilon=var(model))
```

A matrix: 2 × 2 of type dbl

	ilr1	ilr2
ilr1	0.02772303	0.02857064
ilr2	0.02857064	0.05142765

Finally, the parameter estimation variance "F.varEst" is calculated:

In [35]:
```
F.X<-getModelMatrix(model,F.newdata)
F.XX<-kronecker(diag(ncol(predict(model))),F.X)
(F.varEst=F.XX %*% vcov(model) %*% t(F.XX))
```

A matrix: 2 × 2 of type dbl

0.002310252	0.002380887
0.002380887	0.004285638

The steps for calculating the prediction and the parameter estimation variance are repeated for

the V and XXF units:

```
In [36]:  V.newdata<-data.frame(c="V")
          (V.prediction<-ilrInv(predict(model,newdata=V.newdata),orig=Y))
```

```
      gr80     y30-80      y0-30
0.2105087 0.4154923 0.3739990
attr(,"class")
[1] acomp
```

```
In [37]:  V.X<-getModelMatrix(model,V.newdata)
          V.XX<-kronecker(diag(ncol(predict(model))),V.X)
          (V.varEst=V.XX %*% vcov(model) %*% t(V.XX))
```

A matrix: 2 × 2 of type dbl

0.002520275	0.002597331
0.002597331	0.004675241

```
In [38]:  XXF.newdata<-data.frame(c="XXF")
          (XXF.prediction<-ilrInv(predict(model,newdata=XXF.newdata),orig=Y))
```

```
      gr80     y30-80      y0-30
0.3001922 0.3452792 0.3545285
attr(,"class")
[1] acomp
```

```
In [39]:  XXF.X<-getModelMatrix(model,XXF.newdata)
          XXF.XX<-kronecker(diag(ncol(predict(model))),XXF.X)
          (XXF.varEst=XXF.XX %*% vcov(model) %*% t(XXF.XX))
```

A matrix: 2 × 2 of type dbl

0.003080336	0.003174515
0.003174515	0.005714183

Next, the confidence level for the condidence and prediction regions is defined:

```
In [40]:  alpha=0.05
```

Finally, the plot is drawn: First, the grid of the plot is drawn. Next the measured data is plotted. Then, the predictions are plotted. Last, the corresponding confidence ellipses and the prediction ellipses are plotted. They are eliptical in the ilr-space. Hence, the deviation from ellipses results from the back-transformation to the Euclidian space. Finally, the legend is added.

```
In [41]:  # plot preparation grid
          plot(acomp(c(0.2,0.6,0.2)),pch='.',col='grey',cex=0.5,
               labels=c(">80 mm","30-80 mm","0-30 mm"),axes=TRUE)
          myTicks<-1:9/10
          for (i in myTicks){
              x1<-i
              x2<-1:9999/10000*(1-x1)
              x3<-1-x1-x2
              x1<-1-x2-x3
              x<-acomp(cbind(x1,x2,x3))
              plot(x,pch='.',col='grey',add=TRUE)

              x2<-i
              x3<-1:9999/10000*(1-x2)
              x1<-1-x2-x3
              x2<-1-x1-x3
              x<-acomp(cbind(x1,x2,x3))
              plot(x,pch='.',col='grey',add=TRUE)
```

```
    x3<-i
    x1<-1:9999/10000*(1-x3)
    x2<-1-x3-x1
    x3<-1-x1-x2
    x<-acomp(cbind(x1,x2,x3))
    plot(x,pch='.',col='grey',add=TRUE)
}

# plotting
my.lwd<-3
plot(Y[X$c=="F",],pch=6,add=TRUE,col="red",lwd=my.lwd)
plot(Y[X$c=="V",],pch=4,add=TRUE,cex=0.5,col="green",lwd=my.lwd)
plot(Y[X$c=="XXF",],pch=1,add=TRUE,cex=0.5,col="blue",lwd=my.lwd)

plot(F.prediction,pch=17,add=TRUE,col="red",lwd=my.lwd)
plot(V.prediction,pch=15,add=TRUE,col="green",lwd=my.lwd)
plot(XXF.prediction,pch=19,add=TRUE,col="blue",lwd=my.lwd)

ellipses(F.prediction,ilrvar2clr(F.varEst),r=ConfRadius(model,1-alpha),
        col="red",lwd=my.lwd)
ellipses(F.prediction,ilrvar2clr(varEpsilon),r=ConfRadius(model,1-alpha),
        col="red",lty=3,lwd=my.lwd)

ellipses(V.prediction,ilrvar2clr(V.varEst),r=ConfRadius(model,1-alpha),
        col="green",lwd=my.lwd)
ellipses(V.prediction,ilrvar2clr(varEpsilon),r=ConfRadius(model,1-alpha),
        col="green",lty=3,lwd=my.lwd)

ellipses(XXF.prediction,ilrvar2clr(XXF.varEst),r=ConfRadius(model,1-alpha),
        col="blue",lwd=my.lwd)
ellipses(XXF.prediction,ilrvar2clr(varEpsilon),r=ConfRadius(model,1-alpha),
        col="blue",lty="dotted",lwd=my.lwd)

legend(0,1, legend=c("F data","V data", "XXF data", "F prediction", "V prediction",
                    "XXF prediction", "confidence region", "prediction region"),
      col=c("red", "green", "blue", "red", "green", "blue", "black", "black"),
      lty=c(0,0,0,0,0,0,1,3),
      pch=c(6,4,1,17,15,19,NA,NA),cex=0.8)
```
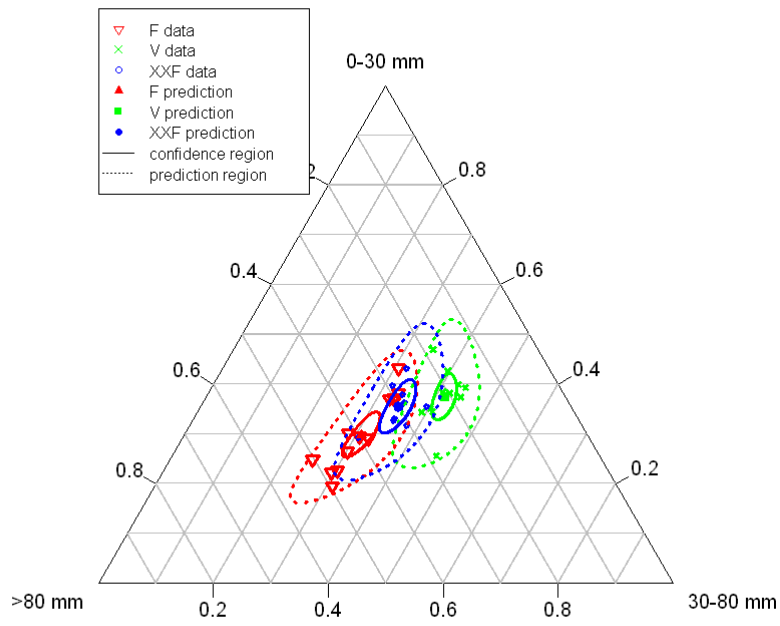
Reference

40. van den Boogaart, K.G.; Tolosana-Delgado, R. *Analyzing Compositional Data with R*; Springer: Dordrecht, 2013, ISBN 978-3-642-36809-7.