

## Selection of optimal conversion path for willow biomass assisted by near infrared spectroscopy

Anna Sandak<sup>(1-2)</sup>,  
Jakub Sandak<sup>(1-2)</sup>,  
Bogusława Waliszewska<sup>(3)</sup>,  
Magdalena Zborowska<sup>(3)</sup>,  
Miroslaw Mleczek<sup>(4)</sup>

Willow (*Salix* sp.) is one of the most common hardwood species suitable for short-rotation coppice. It can be converted to different products, including chemicals, fuels, fibers or furniture. It may also be used in agriculture and environmental engineering. Molecular composition of biomass and its physical properties highly influence effectiveness of its chemical, thermo-chemical or mechanical-chemical conversion. Therefore, it is challenging to provide biomass feedstock with optimized properties, best suited for further downstream conversion. The goal of this research was to establish a procedure for determination of the willow biomass optimal use cultivated in four different plantations in Poland. A special attention has been paid to the application of the near infrared spectroscopy for evaluation of biomass chemical composition and its physical properties. Near infrared spectroscopy (NIR) could be an alternative to standard analytical methods supporting the research and development of biomass production technologies. Partial least squares regression models for quantitative prediction of wood chemical components (lignin, cellulose, hemicellulose, hemicellulose and extractives) and high heating values were developed. The residual prediction deviation (RPD) values confirm the applicability of chemometric models for screening in breeding programmes (for lignin, cellulose and extractives content) and for research in the case of high heating value. The analysis of NIR spectra highlighted several peculiarities in the chemical composition of the investigated willow clones. Finally, a knowledge-based expert system and a prototype automatic NIR system allowing the computation of a “suitability index” based on PLS models and dedicated to selection of optimal biomass conversion path, was developed.

**Keywords:** Willows, NIR Spectroscopy, Optimal Conversion, Biomass Feedstock

### Introduction

Forests cover more than 40% of the area of Europe and are considered as the main supplier of ligno-cellulosic resources. However, some alternative sources for delivering biomass are recently considered as an opportunity, especially when regarding the production of biofuels. Woody energy crops, for instance, are grown solely for the production of renewable energy (Benetka et al. 2014). Fast growing trees or shrub species (usually poplar, willow and

eucalyptus) are cultivated on arable land with an intensive management including fertilizing, irrigation and weed/pest/disease control. The high productivity and the possibility to establish plantations close to the conversion plants make the energy crops a valuable and strategic resource of biomass, especially in highly populated agricultural-urban areas.

Willows (genus *Salix*), belonging to the family *Salicaceae*, are very heterogeneous, comprising over 300 species. Their varia-

tion in size and growth form makes them among the most promising rapidly growing species, considered especially suitable for SRC. Willow possesses high genetic diversity and has great potential for genetic improvement through traditional breeding programmes. Willow yield depends on many factors, such as site, water availability, weed control, planting density, light and temperature.

The use of willow biomass requires knowledge of the achievable yield and its energy value to ensure a regular supply for industry and power plants. The other important issue is the chemical composition of biomass as a feedstock for industry or energy production (Krzyżniak et al. 2014). Willow biomass compared with annual (e.g., wheat straw) and perennial grasses (e.g., *Miscanthus* spp.) has generally a higher content of lignin and energy value, lower cellulose, hemicelluloses and ash content as well as lower concentration of elements such as K, Na, Cl, Si, and S (Karp et al. 2011). These characteristics favor willow biomass for combustion, pyrolysis and gasification instead of biological conversion to biofuels. Willow biomass can be therefore used for the production of primary fuels (in form of chips, briquettes and pellets) and in some cases for secondary

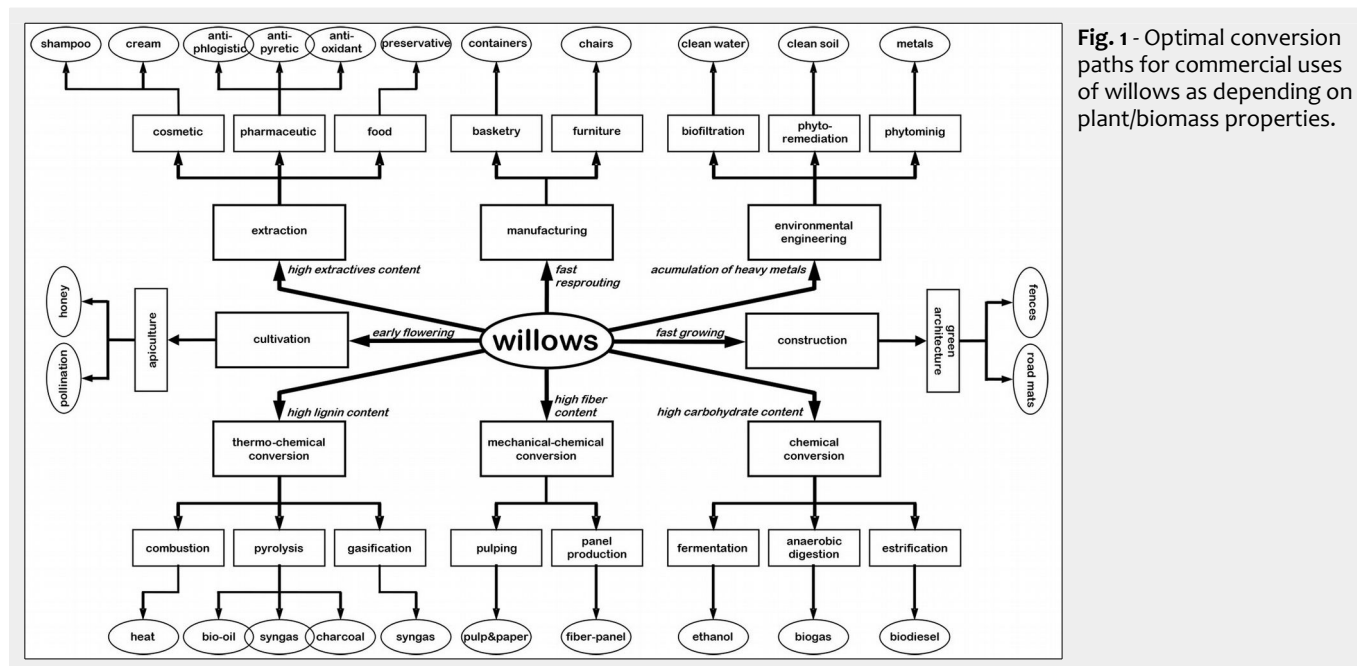
□ (1) CNR-IVALSA, Trees and Timber Institute, via Biasi 75, I-38010 San Michele all'Adige (Italy); (2) University of Primorska, Faculty of Mathematics, Natural Sciences and Information Technology, Glagoljaška 8, 6000 Koper (Slovenia); (3) Poznan University of Life Sciences, Institute of Chemical Wood Technology, ul. Wojska Polskiego 28, 60-637 Poznan (Poland); (4) Poznan University of Life Sciences, Department of Chemistry, ul. Wojska Polskiego 75, 60-625 Poznan (Poland)

@ Jakub Sandak ([sandak@ivalsa.cnr.it](mailto:sandak@ivalsa.cnr.it))

Received: Jan 18, 2016 - Accepted: Feb 13, 2017

**Citation:** Sandak A, Sandak J, Waliszewska B, Zborowska M, Mleczek M (2017). Selection of optimal conversion path for willow biomass assisted by near infrared spectroscopy. *iForest* 10: 506-514. - doi: [10.3832/ifor1987-010](https://doi.org/10.3832/ifor1987-010) [online 2017-04-20]

Communicated by: Giorgio Alberti



**Fig. 1** - Optimal conversion paths for commercial uses of willows as depending on plant/biomass properties.

fuels (bioethanol or wood gas). Willows bark can be the source of salicylic acid, a component used in pharmaceutical, cosmetic and food industries. Willow chips can be also a valuable resource for wood-based panel industry. Apart from economic aspects, cultivation of willows has several ecological advantages, such as the capacity of filtering contaminants from water in wastewater management. An important advantage is willows' capability to accumulate toxins from contaminated sites (phytoremediation). Finally, willow plants can visually improve landscape designs and serve as living snow fences protecting strategic areas from drifting snow (Rowe et al. 2011).

It is required to characterize large populations of plants in order to phenotype biomass to distinguish between genetic and environmental controls over individual bioenergy traits (Tanger et al. 2013). Several crossing strategies have been recently employed within breeding programs in order to increase yields, enhance disease and pests' resistance, improve growth form and resistance to the frost or dry stresses (Tognetti et al. 2013). Genotypes that perform well according to the selection criteria (pest and disease resistances, estimation of yield and chemical composition) are chosen for future plantations (Benetka et al. 2014).

Detection of differences between willow species/clones might be conventionally achieved through the analysis of growth and morphologic characteristics, though this is a time consuming approach. Extensive phenotyping, combined with genetics and genomics, has substantially improved the understanding of the basic traits in willows. It allows targeted breeding *via* marker-assisted selection (Hanley & Karp 2014). Other methods, such as wet chemistry, based on the laboratory testing are mostly

used to study composition of biomass. Disadvantages of such methods are their long measurement time, the requirement of sophisticated laboratory hardware and relatively high costs associated with chemical and/or genetic testing.

Near infrared (NIR) spectroscopy can be an alternative to ordinary analytical methods. NIR was successfully applied to the characterization of woody biomass chemical composition and digestibility (Hou & Li 2011). Several researchers also focused on herbaceous feedstocks (Kelley et al. 2004, Ye et al. 2008). Jensen et al. (2006) compared NIR and dielectric devices for moisture measurement of wood chips from short rotation coppice, among others. Real time prediction of biomass gross calorific value using visible and near infrared spectroscopy was presented by Everard et al. (2012) by analyzing *Miscanthus* and *Salix* species. The attempt for analyzing willows clones by FT-NIR was also reported (Sandak & Sandak 2011).

High-throughput characterization of biomass from bioenergy crops/feedstock is critical for breeding and selection of new cultivars suitable for conversion processes and other industrial practices. The overall aim of the sustainable and economically efficient biomass conversion is to provide a biomass feedstock with optimized properties, best suited for further downstream conversion. Improvements in total biomass yield, biomass quality, and biomass energy content must be matched with the conversion process for which the willow is grown as a feedstock (Serapiglia et al. 2015). Various aspects of the biomass quality could be of interest depending on the desired end-use. A schema of some currently available conversion paths for utilization of willows is presented in Fig. 1. It is noticeable that several possible ways for biomass utilization are available nowadays. However,

it is crucial to correctly define most favorable biomass characteristics in order to correctly select optimal conversion processes.

The report presented here is part of a wider research campaign, where various types of willow clones were cultivated in four plantations in Poland. The overall aim was to select highly productive willow clones that are mostly adapted to particular soil and climate characteristics, being tolerant for disease and heavy metals (Szczukowski et al. 2005). The primary goal of this study was to advance the usage of NIR spectroscopy in evaluating the chemical composition of selected willow species harvested from these experimental plantations. The additional challenge was to develop an original procedure for identification of optimal downstream conversion, best suited for specific biomass feedstock.

## Material and methods

### Experimental samples

Two- and three-year-old willows representing eight different species were collected in Poland on four experimental plantations in Olsztyn, Zielonka, Siemianice and Glogow (see Fig. S1 in Supplementary material). All coppices were cultivated in a similar scheme, regarding plant density and management. No fertilization was applied during plant growth. The only variations among sites were related to the specific climatic and soil characteristics slightly differing in all locations. Mazury Region (Olsztyn plantation, #1) is under the direct influence of the Baltic Sea with high influence of continental climate. Amplitudes of temperatures are large; summer is mild but short in contrast to long and cold winter. It is one of the coldest areas of Poland. Regions of Silesia (Glogow, #4) and Great Poland (Zielonka, #2; Siemianice, #3) are character-

ized by a predominance of oceanic influences. Therefore, the yearly temperature variations are smaller, with long and hot summers but relatively short and mild winters.

Each sample set was in the form of wooden chips (including bark) produced by chipping several plants harvested at the trial fields in February. Chips were mixed on-site and randomly selected 0.5 kg sample package of the fresh mixture was supplied for further laboratory tests. Chips were conditioned in the climatic chamber (20 °C and 60% RH). After reaching the equilibrium moisture content chips were milled with the laboratory mill Pulverisette 15<sup>®</sup> (Fritsch GmbH, Idar-Oberstein, Germany). The final fraction was obtained by sieving and included particles less than 1 mm in size. Species, taxon, cultivation site and age of the investigated samples are summarized in Tab. 1.

### Chemical analysis

Traditional “wet” chemical analyses were performed parallel to the spectroscopic measurements. Analyses included determination of the main chemical components’ content and were conducted according to well established methodologies. Cellulose content was determined according to the Seifert’s method (Browning 1967), using a mixture of acetylacetone, 1,4-dioxane and hydrochloric acid to isolate cellulose. The analysis of the holocellulose content was conducted according to the Chlorite method (Browning 1967), using NaClO<sub>2</sub> as reagent. The determination of acid-insoluble lignin was assessed according to T 222 om-06 standard TAPPI method (TAPPI 2006), using 72 % sulphuric acid to hydrolyze and solubilise carbohydrates. Hemicellulose content was calculated by subtracting cellulose amount from holocellulose. Extractives soluble in hot water were determined according to TAPPI standard (TAPPI 2008). All the chemical analyses were repeated with three replicates for each sample and maximum standard deviation of results within each test was considered as an indicator of the measurement error. The sum of lignin, holocellulose and extractives analysis consistently appear to be greater than 100% in some investigated materials since the results were calculated in relation to wood dry matter. This phenomenon results from imperfection of classical wet methods which are used during determination of concentration of main and minor constituents of ligno-cellulosic materials (Krzyżniak et al. 2014).

High heating values were determined experimentally according to PN-81/G-04513 standard (PN-81/G-04513 1981). Small samples (~1 g) of biomass were completely combusted in the atmosphere of oxygen at 3 MPa in a ZKL-4 calorimeter (Precyzja-Bit, Bydgoszcz, Poland). The gross calorific value (HHV) of solid fuels was determined as a result.

Tab. 1 - Summary of the willow clones investigated in this study.

no	Species	Taxon code	Cultivation site	Age (years)
#1	<i>S. purpurea</i>	7	Zielonka	2
#2	<i>S. × sepulcralis</i> Simonk	17	Zielonka	2
#3	<i>S. purpurea</i>	19/15	Zielonka	2
#4	<i>S. purpurea</i>	26	Zielonka	2
#5	<i>S. purpurea</i>	66/52	Zielonka	2
#6	<i>S. purpurea</i>	125	Zielonka	2
#7	<i>S. viminalis</i>	13/UWM	Olsztyn	3
#8	<i>S. alba</i>	14/UWM	Olsztyn	3
#9	<i>Salix pontederana</i> Willd	15	Olsztyn	3
#10	<i>S. alba</i>	16/UP	Olsztyn	3
#11	<i>Salix × smithiana</i> Willd	43/207/191	Siemianice	2
#12	<i>Salix × smithiana</i> Willd	52/207/191	Siemianice	2
#13	<i>Salix × smithiana</i> Willd	85/207/191	Siemianice	2
#14	<i>S. triandra</i>	85/3/205	Siemianice	2
#15	<i>Salix × smithiana</i> Willd	207/191 (r20b5)	Glogow	2
#16	<i>S. purpurea × triandra × viminalis</i>	46 (r20b5)	Glogow	2
#17	<i>S. purpurea × triandra × viminalis</i>	52/D/46	Siemianice	2

### Heavy metal content

Part of biomass samples was dried at 105 ± 5 °C for 72 h in the laboratory kiln in order to estimate the dry mass. Chips were grounded afterward for 3 minutes in the laboratory cutting boll mill PM200<sup>®</sup> (Retsch GmbH, Haan, Germany) to the fine powder fraction. Three representative subsamples (0.5000 ± 0.0001 g) from each experimental biomass were mineralized in microwave system CEM Mars 5 Xpress<sup>®</sup> (CEM Co., Matthews, NC, USA) by using a solution of 6 mL 65% HNO<sub>3</sub> and 2 mL 30% H<sub>2</sub>O<sub>2</sub> in 55 mL vessels. The microwave mineralization programme was composed of three stages: (i) first stage – 5 min at temperature 120 °C; (ii) second stage – 10 min at 180 °C; (iii) third stage – 10 min at 200 °C.

The solution was filtered after the digestion through 45-mm filters (Qualitative Filter Papers Whatman, Grade 595: 4-7 µm). The supernatant was diluted with deionized water to final volume of 50 mL. For the purpose of this research we considered Cu, Pb and Zn as phytoextraction indicators because they are often present at polluted sites. Total Cu, Pb and Zn content was analyzed by flame atomic absorption spectrometry (FAAS) using a AA Duo – AA280FS/AA280Z spectrometer (Agilent Technologies, Santa Clara, CA, USA). Calibration curves were prepared with five replicates per each metal concentration, by using out-of-stock solution of 1000 mg L<sup>-1</sup> (analytical grade nitrates (V) standard solutions), dissolved in de-ionized water (Milli-Q Academic System – non-TOC). The deuterium background correction was applied in order to minimize the effect of any interference.

### FT-NIR measurement

Experimental samples, the same as used for “wet chemical analysis”, were measured by using Fourier transform near-infrared spectrometer VECTOR 22-N<sup>®</sup> (Bru-

ker Optics GmbH, Fällanden, Switzerland). The instrument was equipped with a standard fibre-optic probe and germanium-diode detector. Thermoplastic resin Spectralon was used as a reference. The spectral resolution of the instrument was 8 cm<sup>-1</sup>, and the spectral range recorded was from 4000 to 12000 cm<sup>-1</sup> (2500 to 833 nm). The spectral wavenumber interval was 3.85 cm<sup>-1</sup> with zero-filling = 2. Each spectrum was computed as an average of 32 successive measurements. Five independent spectra were measured on each milled sample. Signal pre-processing included computation of derivatives, normalization, and smoothing. Derivatives were calculated according to the Savitzky-Golay algorithm.

### Multivariate data analysis

Commercially available (OPUS 7.0, Bruker Optics GmbH; Unscrambler X, CAMO Software AS, Oslo, Norway) and custom (LabView, National Instruments, Austin, TX, USA) software packages were used for data processing and mining. Principal component analysis (PCA), identity test (IT) and partial least squares (PLS) were used for spectra analysis and chemometric models development.

PCA algorithm discovers unique properties of the spectra and separates a set of input data into groups of peculiar similarities. It allows visualization of natural clustering of the data. The measure of the geometrical overlap of such groups is the selectivity index *S*, calculated according to eqn. 1:

$$S = \frac{D}{T_1 + T_2}$$

where *D* is the distance between average spectra, *T*<sub>1</sub> and *T*<sub>2</sub> are threshold values for groups 1 and 2 computed on the basis of the scatter of principle component scores. *S* < 1 indicates that groups are overlapping,



$S = 1$  when two clusters are in contact, and  $S > 1$  in case of two separated sets. This principle can be applied to determination of the unknown spectra's similarity to the reference group and is implemented within the OPUS 7.0 package as an identity test (IT) tool.

The Partial Least Squares (PLS) algorithm was used for computation of regression models linking near infrared spectra and reference values obtained from the chemical analysis. The optimization procedure was applied in order to obtain the best performing models. Several pre-processing methods were implemented into the optimization procedure. The spectra were divided arbitrary into ten optimization regions to allow automatic selection of the spectral ranges the most suitable for regression models. The cross-validation used in this research was the "leave one out" approach. Finally, for each model several fitting parameters were obtained, including the coefficient of determination ( $R^2$ ), the root mean square error of cross-validation (RMSECV), residual prediction deviation (RPD), the bias (the average difference between the estimated and reference values) and the rank (the number of principal components used in the chemometric model).

#### Algorithms for determination of optimal biomass transformation path

Two diverse numerical procedures were developed for quantification of the biomass stock suitability for different transformation processes.

A knowledge-based "if-then-else" expert system assumed that the biomass is "suitable" for a given process by referring relevant biomass property with a threshold level. The value of the threshold and selection of the property of interest is defined by an expert person on the base of specific

process requirements. Only if the given property exceeds the threshold, then the biomass is considered as "yes – suitable", else it is assumed as "no – not suitable".

In the second approach the "suitability index" is computed by means of a dedicated PLS model. NIR spectra representative for two categories of samples, corresponding to the "suitable" and "not-suitable" resources, are used for training the model. The reference values of not-suitable spectra are set to 0, whereas value is 1 for all spectra representing suitable biomass stock. Separate PLS model is required for each transformation scenario, where an expert person selects most characteristic samples representing both "suitable" and "not suitable" resource. Such PLS model can be then used for computation series of suitability indexes, corresponding to different biomass transformation scenarios, for any NIR spectra of unidentified new sample. The output of the model is an integer number, where its value close to 0 indicates low appropriateness of the resource for a given transformation process.

## Results and discussion

#### Alternative paths for conversion of willow biomass

High-throughput characterization of biomass from a bioenergy crop/feedstock is critical for breeding and selecting new cultivars suitable for conversion processes and other industrial practices. The overall aim of the sustainable and economically efficient biomass conversion is to provide a biomass feedstock with optimized properties, best suited for further downstream conversion. Improvements in total biomass yield, biomass quality, and biomass energy content must be matched with the conversion process for which the willow is being grown as a feedstock (Serapiglia et al.

2015). Various aspects of the biomass quality could be of interest depending on the desired end-use and several possible ways for biomass utilization is available nowadays. However, it is crucial to correctly define the most favorable biomass characteristics in order to correctly select optimal conversion processes.

#### Effect of biomass chemical composition on conversion

The overall bio-conversion efficiency of biomass rich in cellulose/hemicellulose is superior comparing to stocks with the high lignin content. This is due to the fact that biodegradability of cellulose is significantly easier and faster than that of lignin.

The chemical composition of the experimental samples used in this research is summarized in Tab. 2. The contents varied between the investigated willow species, especially for holocellulose (from 65.6 to 75.6%) and cellulose (35% to 41.6%). It can be stated, however, that samples #7, #8, #9, #15 and #16 are the most suitable for bio-ethanol production due to the highest holocellulose contents.

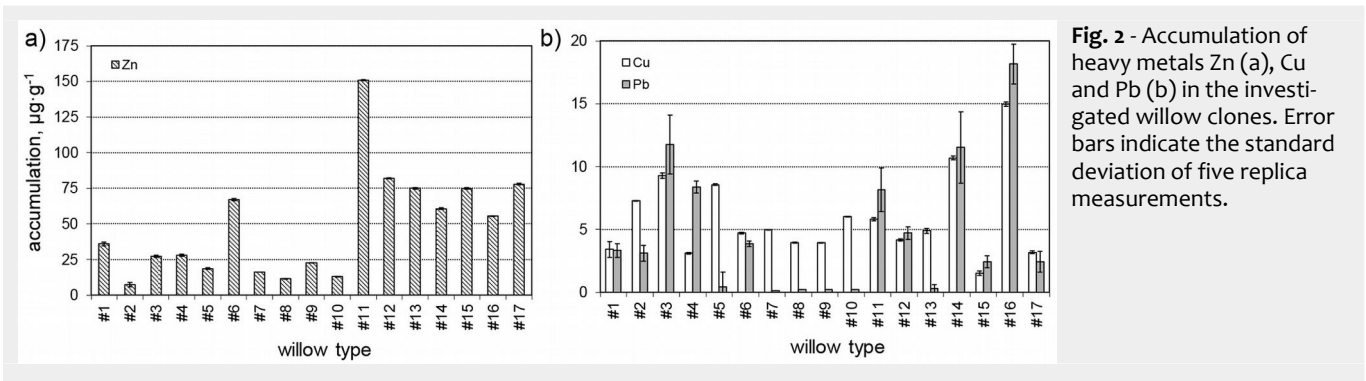
The alternative use of these resources may be related to paper or fiber board production, where low lignin content improves the yield and product quality. On the other hand, the cellulose isolated from willow biomass during pulping process can be post-processed and utilized in production of cellophane, cardboard, paper, or cellulose ethers such as acetate, rayon and nitrates. Cellulose might be also used for manufacturing high-value products: cosmetics, textiles and pharmaceutical drugs. Another derivative such as levulinic acid is a precursor of many pharmaceuticals, plasticizers and a platform for biofuels (Krzyżniak et al. 2014).

Hemicelluloses may be considered as a resource for further chemical transformation into acetic acid, formic acid, furfuryl alcohol, methyltetrahydrofuran and furan (Sinha et al. 2000, Krzyżniak et al. 2014).

Phenolic compounds from lignin degradation slow down enzymatic digestion and fermentation of carbohydrates, and reduce the yield of bio-ethanol. High lignin content is also problematic if the wood is used as a raw material for fibre production, which means that more matter has to be eliminated during the delignification procedure reducing the yield. Conversely, high lignin content can be in some cases considered as an advantage. More lignin increases the heating value of biomass when converted to energy through combustion. The lignin content in the investigated willows ranged from 24% to 28%. For that reason, samples #11, #12 and #17 have the highest potential for conversion into primary fuels. Lignin can be also potentially used as a raw material for manufacturing high-value products, e.g., vanillin, biopolymers in petro-chemistry, pesticides and others (Krzyżniak et al. 2014). Lignin's native structure suggests that it could play an essential role as a new

**Tab. 2** - Chemical composition of the investigated willow clones (in %) and High Heating Value (HHV, MJ kg<sup>-1</sup>). Max  $\sigma$  indicates the maximum standard deviation of results within a given test.

no	Lignin	Cellulose	Holocellulose	Hemicellulose	Extractives	HHV
Max $\sigma$	± 0.55	± 1.00	± 2.70	-	± 0.96	± 0.09
#1	24.50	36.44	70.15	33.71	13.88	20.25
#2	25.01	36.23	65.59	29.36	14.04	20.18
#3	25.07	35.80	68.20	32.40	13.20	20.10
#4	25.24	36.03	70.75	34.73	14.09	20.01
#5	24.22	36.90	69.57	32.67	11.31	20.15
#6	24.52	35.00	67.84	32.84	13.89	19.76
#7	26.92	41.64	74.43	32.76	7.60	19.66
#8	24.98	40.44	70.35	29.96	7.52	19.60
#9	23.95	38.39	70.57	31.84	8.32	19.62
#10	24.53	37.63	68.56	31.77	8.51	19.73
#11	27.13	36.66	72.31	35.60	9.63	19.29
#12	28.20	37.18	72.86	36.00	5.77	19.23
#13	25.81	37.94	73.27	36.00	7.57	19.21
#14	26.42	37.46	71.85	34.70	8.09	19.72
#15	26.22	40.53	75.22	34.40	7.97	20.29
#16	25.58	38.59	70.42	31.60	4.50	19.56
#17	27.00	37.86	75.58	37.98	7.75	19.58



**Fig. 2** - Accumulation of heavy metals Zn (a), Cu and Pb (b) in the investigated willow clones. Error bars indicate the standard deviation of five replica measurements.

chemical feedstock, particularly in the formation of supramolecular materials and aromatic chemicals (Holladay et al. 2007).

Extractives include simple sugars, fats, waxes, gums, terpenes, pectines, resins, fatty acids etc., and are soluble in solvents such as water, ethers and alcohol. The quantity of extractive components varied noticeably between the investigated willows, and in some cases variation reached 9%.

The effect of extractive components on biomass conversion is very complex but in general it is similar to that of lignin, i.e., extractives increase the heating value of biomass but reduce the yield of bio-fuels and fibres. It is due to the fact that the heating value of wood carbohydrates (e.g., holocellulose) is ~16 MJ kg<sup>-1</sup>, while lignin is ~25 MJ kg<sup>-1</sup> and some extractives can reach ~32 MJ kg<sup>-1</sup> (Sandak & Sandak 2011).

High content of extractives (glycosides and tannins) is required when raw resources are used within pharmaceutical industry. Willows #1 to #6 might be utilized in such a way since extractives amount is significantly higher than in the rest of the tested biomass.

**Energetic use of willow**

The heating value of wood is directly related to its chemical composition and indicates the quantity of thermal energy stored in the fuel per unit mass. High heat values (HHV) of the investigated willows

varied in the range from 19.2 to 20.3 MJ kg<sup>-1</sup> (Tab. 2). The most suitable clones for heat generation were: #1, #2, #3, #4, #5 and #15, originated from Zielonka or Glogow.

The reliable and straightforward estimation of HHV is of a great interest for biomass producers and users. It is possible to estimate this value by knowing the chemical composition of biomass and applying mathematical formula. It has to be also mentioned, however, that such a method is far from optimal as necessary reference values have to be obtained by a costly and time-consuming wet chemistry analysis. Hence, an alternative and simplified method (based on NIR) has been evaluated in this research as described below.

**Phytoremediation with willow biomass**

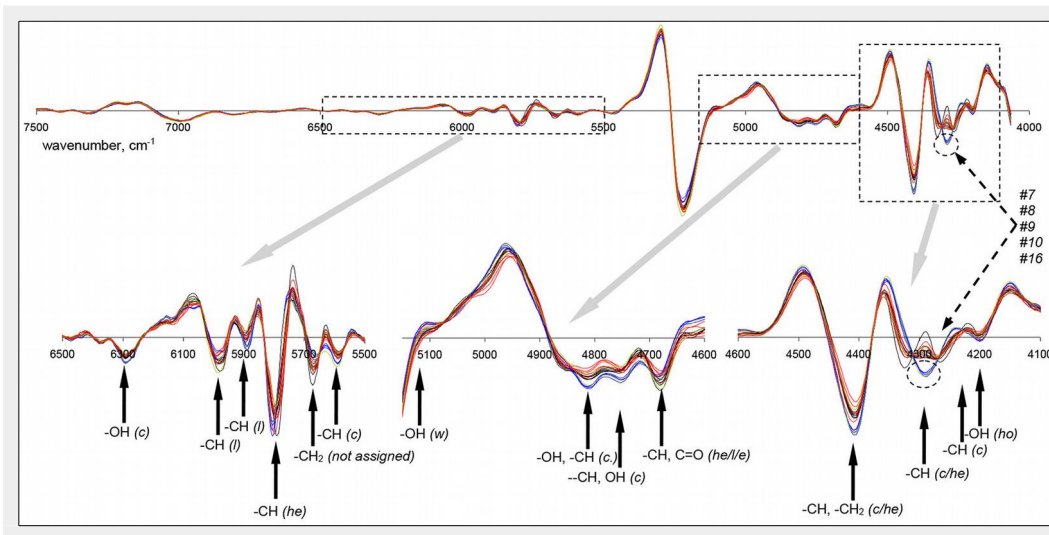
The ability for fast propagation of plants with simultaneous accumulation of heavy metals is another essential characteristic to be considered while optimizing a conversion path. Perennial nature, fast growth, and ability to survive in relatively harsh soil conditions are the main reasons of using willows for phytoremediation (Pulford & Watson 2003). Two strategies for utilization of plants capable of heavy metals accumulation can be applied; the pollutant removal from contaminated soil by the plant root system, and/or phytomining of heavy metals from the impure air by means of the assimilation system of the living plants (Rascioa & Navari-Izzob 2011).

Results from FAAS as related to content of Zn, Cu and Pb are summarized in Fig. 2. High accumulation of heavy metals (especially zinc) was observed in *Salix viminalis* (samples #11, #12, #13, #14, #15, #16 and #17). All of these were cultivated in Siemianice and Glogow plantations. The highest accumulation of Zn (~150 ug g<sup>-1</sup>) was noticed for willow #11. Sample #15 was considered by experts as a very valuable taxon, since it was the fastest growing among others in addition to ability for heavy metals accumulation. *Salix pondeterana* (#9) was identified as the most rapidly propagating among all the investigated clones. Fast colonization, growing and fostering make this particular willow an effective species to be pioneer flora planted on young or recovered sites.

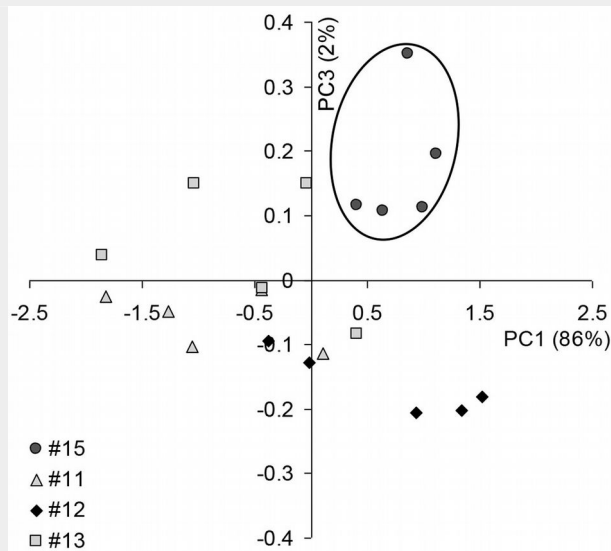
Willows #3, #4, #11, #14, and #16 are considered as the most suitable for accumulation of lead. Copper was mainly absorbed by willows #3, #5, #14 and #16, where samples #3, #14 and #16 were able to accumulate both Cu and Pb. It was found that all clones cultivated in Olsztyn possessed the lowest prospective as phytoremediators.

**Characterization of willow biomass with near infrared spectroscopy**

Averaged second derivative NIR spectra of the investigated samples are presented in Fig. 3. All curves have, in general, very similar shapes. However, clear variations may be observed in some bands. Several



**Fig. 3** - NIR spectra for experimental willows of different clones and provenances. The chemical components associated to the specific bands are: (l) lignin; (c) cellulose; (he) hemicellulose; (ho) holocellulose; (e) extractives; (w) water.



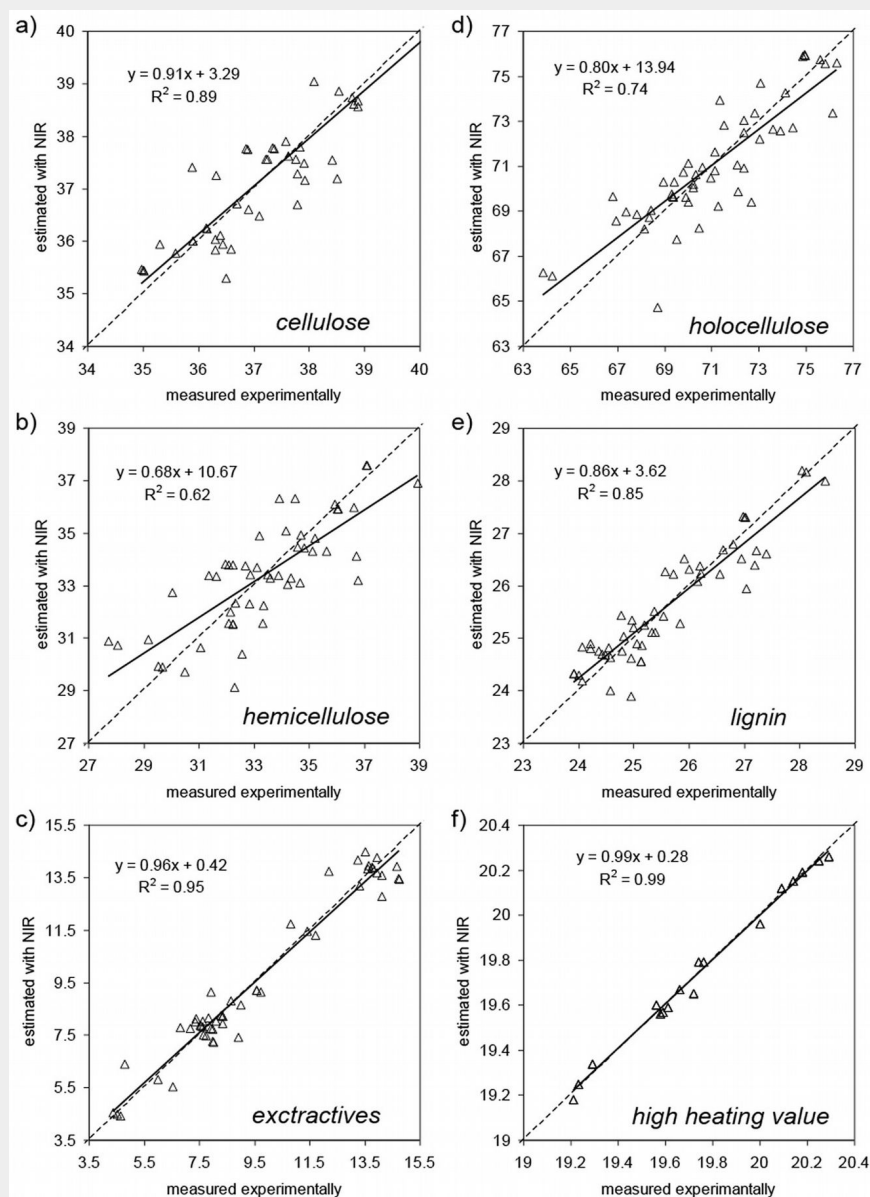
**Fig. 4** - PCA of *Salix x smithiana* Willd clones cultivated in different plantations based on raw spectra in the range 12000-4000  $\text{cm}^{-1}$ .

peaks have more intense absorbance especially in the case of willows originated from Olsztyn (#7, #8, #9, #10) and Glogow (#16). Opposite trends can be noticed in band 4286  $\text{cm}^{-1}$  related to C-H stretching and C-H deformation in hemicelluloses. Peak 4404  $\text{cm}^{-1}$  ( $\text{CH}_2$ ), range 4560-4850  $\text{cm}^{-1}$  related to carbohydrates (both cellulose and hemicellulose) and peaks 5464  $\text{cm}^{-1}$  and 5587  $\text{cm}^{-1}$  related to O-H and C-H bonds in crystalline and semi-crystalline region of cellulose are also different in regard to these samples. A spectral signal attributed to hemicelluloses (furanose and pyranose) at 5800  $\text{cm}^{-1}$  is shifted to 5805  $\text{cm}^{-1}$ . Peaks 5890  $\text{cm}^{-1}$  (aromatic C-H due to lignin), 6295  $\text{cm}^{-1}$  and 7315  $\text{cm}^{-1}$  related to hydroxyl groups of cellulose are also more intense. Such an observation was confirmed by PCA where a single cluster containing the above willows was clearly separated from the others (Fig. 4). In fact, according to the studies of Serapiglia et al. (2013) and Sandak et al. (2015), the variation in cellulose content within the stem is controlled by environmental factors, among others. The average temperature during the vegetation season 2009-2012 ranged from 14.6 to 15.1  $^{\circ}\text{C}$  in Olsztyn. Total rainfall during this period ranged from 384 to 511 mm, even if precipitation during the growing season was unevenly distributed.

Clustering of NIR spectra might be also explained by the fact that willows from Olsztyn were harvested after 3 years of cultivation period. Szczukowski et al. (2005) reported that cellulose content raises with the increase of the crop rotation time. Rowell et al. (1997) observed that biomass harvested from young woody plants contained less holocellulose and lignin than that from mature plants. Conversely, not all willows cultivated in Olsztyn for three years possessed the highest cellulose content. Thus, the variation observed in NIR spectra is related not merely to the quantity of components, but it is also connected to the quality of constitutive woody polymers. Such variations within the molecular structure of both carbohydrates and lignin in *Picea abies* of a different origin were observed by the authors (Santoni et al. 2015).

#### Qualitative assessment of willow biomass with NIR

Differences in NIR spectra between individuals of the same taxon (207/191) but harvested at different sites were observed in *Salix x smithiana* Willd by means of PCA. Clustering of willows cultivated in Glogow (#15) and three field trials in Siemianice (#11, #12, #13) is presented in Fig. 4. The highest distance (selectivity index) was noticed between sample #15 (grown in Glogow) and all the others (see Tab. S1 in Supplementary material). Willow #15 possesses relatively high cellulose content and HHV apart from capabilities for Zn accumulation and relatively high hemicellulose content (as similar to the other samples from this



**Fig. 5** - PLS calibration models for cellulose (a), hemicellulose (b), extractives (c), holocellulose (d), lignin content (e) and high heating value (f) determined for all investigated willow clones.



taxon). Therefore, it is clear that chemical composition and traits of even the same taxons might differ considerably depending on the plantation site and environmental conditions.

### NIR quantitative analysis of willow biomass properties

Partial Least Squares (PLS) algorithms are frequently applied to computation of regression models linking near infrared spectra of matter and reference values related to its properties. Prediction of cellulose, holocellulose, hemicellulose, lignin and extractive contents of the investigated samples were implemented into the Quant2 software module (part of OPUS package). The resulting chemometric models are presented in Fig. 5. PLS models are suitable for screening if  $RPD \geq 3$ , for quality control if  $RPD \geq 5.0$ , and for applied research if  $RPD \geq 8.0$  (Manley 2014).  $RPD$  values obtained in this study indicate that models might be useful for screening in breeding programs while predicting cellulose, lignin and extractive contents (Tab. 3). The dedicated PLS model was also developed for the estimation of the high heating value. Very high  $RPD$  value (9.87) calculated for  $HHV_{NIR}$  proves that this approach might be used for quality control of biomass. It is a great alternative to lab-intensive methodologies, as there is no need for intensive sample preparation and/or the chemical analysis while estimating biomass properties with NIR. We are aware that the reliability of the presented models is limited due to a rela-

**Tab. 3** - Validating parameters for PLS regression prediction models developed for the investigated willow clones. ( $R^2$ ): coefficient of determination; ( $RMSECV$ ): root mean square error of cross-validation; ( $RPD$ ): residual prediction deviation; ( $MSC$ ): multiplicative scatter correction; ( $HHV$ ): high heating value.

Parameter	Cellulose	Holo-cellulose	Hemi-cellulose	Lignin	Extractives	HHV
$R^2$	0.89	0.74	0.61	0.86	0.95	0.99
$RMSECV$	0.588	1.45	1.45	0.453	0.687	0.0343
$RPD$	3.05	1.95	1.6	2.63	4.47	9.87
bias	0.028	-0.0248	-0.0296	-0.00896	0.00402	-0.00137
range ( $cm^{-1}$ )	8450-7498	9403-7491 6401-4246	7501-6098	9403-5450	9403-7498	9403-7424
pre-processing	min-max normalization	1 <sup>st</sup> derivative + MSC	straight line substitution	2 <sup>nd</sup> derivative	2 <sup>nd</sup> derivative	2 <sup>nd</sup> derivative
rank	7	7	6	6	5	5

tively low number of available reference data. However, the data base is continuously updated with ongoing research as well as chemometric models.

### Optimization of the willow biomass transformation path

#### Knowledge-based expert system

Significant variations in chemical content, and consequently in heating values, have been noticed among the investigated clones from four plantations in Poland. It is possible to advise the most appropriate conversion ways, based on statistical analysis and expert evaluation. An original procedure, knowledge-based expert system, was developed for the needs of this re-

search. The algorithm is based on the assumption that the quantity of a determined component has to exceed a certain threshold to consider the given stock as "suitable". The threshold values were set arbitrarily by comparing the researched biomass and available characteristics provided by analytical methods. A summary of the expert system recommendation is presented in Tab. 4, where the suitability of each willow clone to various processes was identified and simply defined as "yes" or "no". It is clear that such straightforward approach should be considered only as a support in the decision making in real operations. As discussed before, there are different feedstock requirements for chemical vs. thermo-chemical conversion path-

**Tab. 4** - Summary of investigated willow clones with the suggested utilization path assessed by expert based on the chemical properties ("yes" or "no") and estimated with near infrared spectroscopy (0 - low suitability, 1 - high suitability). ( $HHV$ ): high heating value; ( $X_c$ ): content of cellulose; ( $X_{he}$ ): content of hemicellulose; ( $X_e$ ): content of extractive components; ( $X_{Zn}$ ): content of zinc; ( $X_{Pb}$ ): content of lead; ( $X_{Cu}$ ): content of copper. The average prediction error is the difference between the expert assessment (no = 0 or yes = 1) and the NIR estimated value of the suitability index.

Sample no.	Suggested utilization path						
	thermo-chemical	mechano-chemical	chemical	pharmaceutical	phytoremediation		
#1	yes / 0.93	no / 0.07	no / 0.13	yes / 0.96	no / 0.19	no / -0.10	no / 0.00
#2	yes / 0.98	no / 0.04	no / -0.01	yes / 0.98	no / 0.44	no / -0.03	no / 0.02
#3	yes / 0.97	no / 0.13	no / 0.01	yes / 1.02	no / 0.57	yes / 0.88	yes / 0.33
#4	yes / 1.05	no / -0.07	yes / 1.04	yes / 0.99	no / 0.61	yes / 0.99	no / 0.20
#5	yes / 0.89	no / 0.13	no / 0.19	yes / 0.80	no / 0.57	no / -0.06	yes / 0.19
#6	no / 0.05	no / -0.05	no / -0.13	yes / 1.14	yes / 0.68	no / 0.15	no / 0.42
#7	no / -0.08	yes / 0.78	no / 0.02	no / 0.03	no / 0.33	no / -0.07	no / 0.50
#8	no / 0.09	yes / 1.10	no / -0.07	no / -0.11	no / 0.06	no / 0.00	no / -0.03
#9	no / -0.05	yes / 0.90	no / -0.06	no / 0.02	no / 0.28	no / 0.15	no / 0.30
#10	no / -0.03	no / 0.04	no / 0.07	no / 0.03	no / 0.09	no / 0.14	no / -0.11
#11	no / -0.02	no / -0.05	yes / 1.01	no / -0.06	yes / 1.19	yes / 0.84	no / 0.31
#12	no / -0.04	no / -0.07	yes / 1.05	no / -0.04	yes / 1.32	no / -0.15	no / 0.22
#13	no / -0.02	no / -0.04	yes / 0.93	no / 0.08	yes / 1.22	no / 0.32	no / 0.35
#14	no / -0.04	no / -0.04	yes / 0.92	no / 0.10	yes / 1.00	yes / 0.86	yes / 0.66
#15	yes / 1.11	yes / 0.95	yes / 0.84	no / 0.00	yes / 1.04	no / 0.09	no / 0.11
#16	no / 0.16	yes / 1.12	no / -0.02	no / 0.06	yes / 1.31	yes / 1.12	yes / 0.60
#17	no / 0.06	no / 0.06	yes / 1.07	no / 0.01	yes / 1.10	no / -0.14	no / -0.05
Discrimination rule	high $HHV$	high content of:			high accumulation of:		
		cellulose	hemicellulose	extractives	Zn	Pb	Cu
Threshold rule	$HHV > 20.0$	$X_c > 38\%$	$X_{he} > 34\%$	$X_e > 11\%$	$X_{Zn} > 55ppb$	$X_{Pb} > 8ppb$	$X_{Cu} > 8ppb$
Average prediction error	0.06	0.08	0.07	0.06	0.29	0.11	0.29

ways. Likewise, the environmental factors are highly influencing the biomass usage. Both, location and seasonal factors, beside the chemical composition should be considered when strategic decisions are taken.

#### NIR suitability index

The expert system approach for the optimization of the biomass transformation path can only be applied after performing a time-consuming and costly chemical analysis. The alternative automatic system has been developed within the frame of this research by means of NIR spectroscopy. PLS models were designed, where the variable to be predicted was a "suitability index" ranging from 0 (not suitable) to 1 (suitable). Separate indexes were computed for each conversion process, including thermo-chemical, mechano-chemical, chemical, pharmaceutical and phytoremediation paths. The test set validation was applied, where independent spectra measured on the analyzed clones were used for prediction. The results of automatic NIR system rating are summarized in Tab. 4 in a form of numerical indexes. Values close to "0" indicate low suitability for the conversion process, while a high value close to "1" suggests the high appropriateness of the biomass for the specific transformation. The number of factors (principal components) used for PLS was three, thus ensuring a relative generalization of models and avoiding over-fitting. The results of NIR prediction correspond closely to that of the expert evaluation. The most inaccurate predictions are on the phytoremediation capacity. It is not surprising as near infrared spectroscopy is not capable of directly measuring heavy metals but only the effect of their presence within the polymeric matrix of wood. Such a system permits rapid, accurate and very low cost possibility for biomass quality control and immediate determination of the optimal conversion path.

We are aware that the initial chemical composition of the biomass is only one of the key parameters affecting the selection of the biomass valuation-utilization chain. In particular, the enzyme hydrolysis rate and ultimate sugar yield are both impacted by higher total sugar content. Similarly, the reactivity of biomass in thermochemical processes, as well as the quality and value of the resulting bio-oil or syngas depend not only on the lignin content or HHV. Therefore, experimental validation of the biomass reactivity and/or product yield/quality in combination with refining NIR chemometric models is indispensable.

#### Conclusions

Seventeen samples of willow originated from four experimental stations in Poland were investigated and characterized by various analytical methods. It was shown that biomass properties vary due to provenance, growth conditions and genetic features. A reliable, rapid and accurate tool

suitable for the determination of biomass chemical composition is necessary to guarantee the supply of feedstock with optimized properties and best suited for specific downstream conversion. It was shown that NIR spectroscopy can be such an alternative technique to the standard analytical methods.

The near infrared technique is sensitive enough for the detection of even minor differences in the molecular structure of wood, as related to the biomass genetics and environment. A prototype automatic NIR system devoted to selection of the optimal biomass conversion path was developed in the frame of this study. The result generated by the system is a "suitability index" ranging from 0 to 1. It indicates the advantage of using certain stock for defined applications. Models were developed separately for varying conversion processes, including thermo-chemical, mechano-chemical, chemical, pharmaceutical and phytoremediation paths. Results of the NIR-based approach were comparable with the expert assessment.

#### Acknowledgments

This research was supported by the National Centre for Research and Development within the framework of the development grant no. NRNN 309708740. Part of this work has been conducted within the framework of the project SLOPE funded by the EU 7<sup>th</sup> Framework Programme for research, technological development and demonstration under the NMP.2013.3.0-2 (grant no. 604129).

#### References

- Benetka V, Novotná K, Stochlová P (2014). Biomass production of *Populus nigra* L. clones grown in short rotation coppice systems in three different environments over four rotations. iForest 7: 233-239. - doi: [10.3832/ifor1162-007](https://doi.org/10.3832/ifor1162-007)
- Browning BL (1967). The chemistry of wood. Interscience Publisher, New York, USA, pp. 407.
- Everard CD, McDonnell KP, Fagan CC (2012). Prediction of biomass gross calorific values using visible and near infrared spectroscopy. Biomass and Bioenergy 45: 203-211. - doi: [10.1016/j.biombioe.2012.06.007](https://doi.org/10.1016/j.biombioe.2012.06.007)
- Hanley SJ, Karp A (2014). Genetic strategies for dissecting complex traits in biomass willows (*Salix* spp.). Tree Physiology 34 (11): 1167-80. - doi: [10.1093/treephys/tpto89](https://doi.org/10.1093/treephys/tpto89)
- Holladay JE, White JF, Bozell JJ, Johnson D (2007). Top value-added chemicals from biomass - Volume II. Results of screening for potential candidates from biorefinery lignin. PNNL-16983, Pacific Northwest National Laboratory, Richland, WA, pp89. [online] URL: <http://www.energy.gov/sites/prod/files/2014/03/f14/pnnl-16983.pdf>
- Hou S, Li L (2011). Rapid characterization of woody biomass digestibility and chemical composition using near-infrared spectroscopy. Journal of Integrative Plant Biology 53 (2): 166-75. - doi: [10.1111/j.1744-7909.2010.01003.x](https://doi.org/10.1111/j.1744-7909.2010.01003.x)
- Jensen PD, Hartmann H, Bohm T, Temmerman

M, Rabierc F, Morsing M (2006). Moisture content determination in solid biofuels by dielectric and NIR reflection methods. Biomass and Bioenergy 30 (11): 935-943. - doi: [10.1016/j.biombioe.2006.06.005](https://doi.org/10.1016/j.biombioe.2006.06.005)

Karp A, Hanley SJ, Trybush SO, Macalpine W, Pei M, Shield I (2011). Genetic improvement of willow for bioenergy and biofuels. Journal of Integrative Plant Biology 53 (2): 151-165. - doi: [10.1111/j.1744-7909.2010.01015.x](https://doi.org/10.1111/j.1744-7909.2010.01015.x)

Kelley SS, Rowell RM, Davis M, Jurich CK, Ibach R (2004). Rapid analysis of the chemical composition of agricultural fibers using near infrared spectroscopy and pyrolysis molecular beam mass spectrometry. Biomass and Bioenergy 27 (1): 77-88. - doi: [10.1016/j.biombioe.2003.11.005](https://doi.org/10.1016/j.biombioe.2003.11.005)

Krzyżniak M, Stolarski M, Waliszewska B, Szczukowski S, Tworowski J, Zaluski D, Snieg M (2014). Willow biomass as feedstock for an integrated multi-product biorefinery. Industrial Crops and Products 58: 230-237. - doi: [10.1016/j.indcrop.2014.04.033](https://doi.org/10.1016/j.indcrop.2014.04.033)

Manley M (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chemical Society Reviews 43: 8200-8214. - doi: [10.1039/C4CS00062E](https://doi.org/10.1039/C4CS00062E)

PN-81/G-04513 (1981). Paliwa stałe Oznaczenie ciepła spalania i obliczanie wartości opalowej [Solid fuels determination of gross calorific value and calculation of net calorific value]. Polski Komitet Normalizacyjny, Warsaw, Poland, pp. 11. [in Polish]

Pulford ID, Watson C (2003). Phytoremediation of heavy metal-contaminated land by trees - a review. Environment International 29 (4): 529-540. - doi: [10.1016/S0160-4120\(02\)00152-6](https://doi.org/10.1016/S0160-4120(02)00152-6)

Rascioa N, Navari-Izzob F (2011). Heavy metal hyperaccumulating plants: how and why do they do it? And what makes them so interesting? Plant Science 180 (2): 169-181. - doi: [10.1016/j.plantsci.2010.08.016](https://doi.org/10.1016/j.plantsci.2010.08.016)

Rowe RL, Hanley ME, Goulson D, Clarke DJ, Doncaster CP, Taylor G (2011). Potential benefits of commercial willow Short Rotation Coppice (SRC) for farm-scale plant and invertebrate communities in the agri-environment. Biomass and Bioenergy 35 (1): 325-336. - doi: [10.1016/j.biombioe.2010.08.046](https://doi.org/10.1016/j.biombioe.2010.08.046)

Rowell RM, Han JS, Bisen SS (1997). Changes in fiber properties during the growing season. In: "Paper and Composite from Agro-based Resources" (Rowell RM, Young RA, Rowell JK eds). Lewis Publishers, Boca Raton/New York/London/Tokyo, pp. 23-37.

Sandak J, Sandak A (2011). FT-NIR assessment of biomass composition of shrub willow clones (*Salix* sp.) for optimal bio-conversion process. Journal of Near Infrared Spectroscopy 19 (5): 309-318. - doi: [10.1255/jnirs.950](https://doi.org/10.1255/jnirs.950)

Sandak J, Sandak A, Cantini C, Autino A (2015). Differences in wood properties of *Picea abies* L. Karst. in relation to site of provenance and population genetics. Holzforshung 69 (4): 385-397. - doi: [10.1515/hf-2014-0061](https://doi.org/10.1515/hf-2014-0061)

Santoni I, Callone E, Sandak A, Sandak J, Dirè S (2015). Solid state NMR and IR characterization of wood polymer structure in relation to tree provenance. Carbohydrate Polymers 117: 710-721. - doi: [10.1016/j.carbpol.2014.10.057](https://doi.org/10.1016/j.carbpol.2014.10.057)

Serapiglia MJ, Cameron KD, Stipanovic AJ, Abra-



- hamson LP, Volk TA, Smart LB (2013). Yield and woody biomass traits of novel shrub willow hybrids at two contrasting sites. *Bioenergy Research* 6 (2): 533-546. - doi: [10.1007/s12155-012-9272-5](https://doi.org/10.1007/s12155-012-9272-5)
- Serapiglia MJ, Gouker FE, Hart JF, Unda F, Mansfield SD, Stipanovic AJ (2015). Ploidy level affects important biomass traits of novel shrub willow (*Salix*) hybrids. *Bioenergy Research* 8: 259-269. - doi: [10.1007/s12155-014-9521-x](https://doi.org/10.1007/s12155-014-9521-x)
- Sinha S, Jhalani A, Ravi MR, Ray A (2000). Modeling of pyrolysis in wood: a review. *Journal of the Solar Energy Society of India* 10 (1): 41-62.
- Szczukowski S, Stolarski M, Tworkowski J, Przyborowski J, Klasa A (2005). Productivity of willow coppice plants grown in short rotations. *Plant, Soil and Environment* 9: 423-430. [online] URL: <http://agriculturejournals.cz/publicFiles/50990.pdf>
- Tanger P, Field JL, Jahn CE, DeFoort MW, Leach JE (2013). Biomass for thermochemical conversion: targets and challenges. *Frontiers in Plant Science* 4 (218): 1-20. - doi: [10.3389/fpls.2013.00218](https://doi.org/10.3389/fpls.2013.00218)
- TAPPI (2006). Acid Insoluble lignin in wood and pulp. T 222 om-06. Technical Association of the Pulp and Paper Industry, New York, USA, pp. 5.
- TAPPI (2008). Solvent extractives of wood and pulp. T 207 cm-08. Technical Association of the Pulp and Paper Industry, New York, USA, pp 4.
- Tognetti R, Cocozza C, Marchetti M (2013). Shaping the multifunctional tree: the use of Salicaceae in environmental restoration. *iForest* 6: 37-47. - doi: [10.3832/ifor0920-006](https://doi.org/10.3832/ifor0920-006)
- Ye XP, Liu L, Hayes D, Womac A, Kunlun H, Sokhansanj S (2008). Fast classification and compositional analysis of cornstover fractions using Fourier transform near-infrared techniques. *Bioresource Technology* 99 (15): 7323-7332. - doi: [10.1016/j.biortech.2007.12.063](https://doi.org/10.1016/j.biortech.2007.12.063)

## Supplementary Material

**Fig. S1** - Localization of plantation trials in Poland: 1 - Olsztyn, 2 - Zielonka, 3 - Glogow, 4 - Siemianice.

**Tab. S1** - Selectivity index calculated for *Salix × smithiana* Willd (taxon code 207-191) for willows #11, #12, #13 and #15.

**Link:** [Sandak\\_1987@suppl001.pdf](mailto:Sandak_1987@suppl001.pdf)