

## Image - Rice Grain Scanner: a three-dimensional fully automated assessment of grain size and quality traits

Rubens Marschalek<sup>1</sup>, Mauricio Cesar Silva<sup>1</sup>, Samuel Batista dos Santos<sup>1</sup>, Johnny Ricardo Manke<sup>2</sup>, Carlos Biegging<sup>2</sup>, Geovani Porto<sup>1</sup>, Ester Wickert<sup>1</sup> and Alexander de Andrade<sup>1</sup>

**Abstract:** *The Image is a scanner developed as a grain classifier for quality control at the rice industry based on Brazilian official norms. It orders the dehulled grains ensuring that each grain would pass individually, in free fall, while the grain is analysed from different sides, covering its whole surface. It ensures a precise three-dimensional measurement of grain size, chalkiness, defects of the grain, milling quality, given out a total of 39 traits/classes/defects/values, which are sent to a excel Microsoft spreadsheet. This is managed through a digital platform, which analysis routine and layout were developed and designed by Selgron and Epagri to fit the needs of research. The scanner and its software reach outputs that enhance rice breeding efficiency for grain quality, performing it faster, precisely and with a high-throughput phenotyping than ever before, especially interesting in very early breeding generations.*

**Key words:** *Software, Oryza sativa, breeding for grain quality, high-throughput phenotyping, milling quality.*

### INTRODUCTION

Most of rice breeding programs have, for many years, stressed out yield as the main trait to be observed during the development of new lines. Nevertheless, and to fit the different needs or desires of consumers across the world, grain quality is also evaluated and used for selection in breeding programs. This is important from the point of view of the consumer, even to the rice industry itself, since quality patterns vary along the kind of consumers worldwide, and even the patterns are changing, because consumers desire are moving forward. To worsen, and despite its importance, rice quality traits are not easy to measure, and quite often are evaluated as qualitative instead of quantitative ones.

Besides, even traits as grain size, directly related to grain quality or marked, typically a quantitative trait, are too much depended on sample size. For example, if it is necessary to determine de grain size of a new variety, generally only few grains are used to be measured to get the information, which leads to a basic sampling error, which, on the other hand, will lead to a wrong estimation of average grain size. At Epagri's Rice Breeding Program, for example, a sample of only ten grains was usually taken from a new rice line in early generation stages to estimate the three dimensions measures by using a pachymeter. This low number of grains implies in a clear sampling error. Other breeding programs

Crop Breeding and Applied Biotechnology  
17: 89-97, 2017  
Brazilian Society of Plant Breeding.  
Printed in Brazil  
<http://dx.doi.org/10.1590/1984-70332017v17n1s15>

**\*Corresponding author:**  
E-mail: [rubensm@epagri.sc.gov.br](mailto:rubensm@epagri.sc.gov.br)

**Received:** 15 December 2016  
**Accepted:** 20 December 2016

<sup>1</sup> Epagri - Santa Catarina State Agricultural Research and Rural Extension Agency Itajaí Experiment Station, Rodovia Antônio Heil 6800, 88.318-112, Itajaí, SC, Brazil

<sup>2</sup> Selgron Industrial Ltda, Rua Hermann Althoff, 220, 89.066-355, Blumenau, SC, Brazil

use to measure 100 to 1000 grains to evaluate a rice genotype, but this is usually done up to  $F_5$  breeding generations. Such evaluations efforts are usually difficult to proceed in the usually large number of lines in breeding programs.

If chalkiness is the trait to be evaluated it became even more difficult to get good evaluations, since it is treated as a usually qualitative trait estimated by a grade observed by a trained technician. Another problem is that chalkiness is irregularly distributed, however some patterns could be observed between different genotypes. Sometimes chalkiness affects a small part of many grains, other times lead to complete chalked grains, but only a few ones are affected. So, it is not only important to know how many grains have chalked areas, but how large is this affected area throughout the whole grains in a sample of rice.

The solutions available for rice scientists to handle the selection of such complex traits are some software that, working together with images took usually by an ordinary desktop scanner, measure and/or analyse grains in two dimensions. This is found at the 'SmartGrain' (Tanabata et al. 2012), designed for breeding efforts, or the Rice/Grain Analyzer Software 6980 (Advanced Version) (Osaw Industrial Products Pvt. Ltd., Indian). There are even some solutions that bring together engines and software, to do it also in one way, integrating the scanning and analysis processes, like the Classifier S21 – Rice Statistical Analiser (Máquinas Suzuki S.A., Brazil), the RN300 Rice Quality Analyser (Kett, US), the Satake RSQ10A Grain Scanner (Satake, Australia), or the SeedCount SC5000 Rice Analyser (Next Instruments Pty Ltd, Australia). Recent approaches given by the GrainScan free software (Whan et al. 2014) combine elements of 'SmartGrain' and 'Seedcount'. Such solutions seem to be the updated possibilities of breeding for cereal quality worldwide, but not always are specific for rice.

Since Epagri's Rice Research Team were presented to the 'Image Classifier', developed by Selgron (Blumenau-SC-Brazil [www.selgron.com.br](http://www.selgron.com.br)), it became clear to the breeders that this instrument, supported by an appropriate software, offers a unique chance to enhance the way of collecting grain quality data in a more precise and quickly way than even before. The 'Image Classifier' was developed to support the rice industry by quickly analyzing samples of about 3300 milled grains (100g) in twelve minutes (according Selgron). The industrial software verifies how each sample match to the different rice type classifications of the Agriculture Ministry of Brazil at its norms (MAPA, 'Instrução Normativa nº 06, Feb 16<sup>th</sup>, 2009'). In this way, it is a very useful equipment to manage the quality control at the rice industry.

However, like predicted by Epagri's team, for the research use of this scanner, changes, adaptations and the way to organize the rough data were necessary. The software platform to manage samples at the Image Rice Grain Scanner was so, developed by Selgron under Epagri's scientists demand, advisement and supervision, so that it ensures a full use directly applied to the breeding/research needs. This includes some routines, like sample identification through a bar code reader, step by step protocol for analysis, data layout at the spreadsheet, and even new approaches, like the introduction of standard deviation at size measurements, total sample grain chalked area (besides percentage of chalked grains), and milling quality, the last one based on the whole and broken grains estimated weight. The final version of the Image Research Software Platform was upgraded at May 2016.

## HARDWARE

The device has it Patent required under number PI 1001551-5 de 20.05.2010, as 'Equipamento classificador de grãos para estabelecimento de qualidade'. The equipment (100W, 220V, 50-60Hz) is operated by means of a microcomputer integrated with the Image scanner (called classifier in the rice industry version). The Image consists of the scanner itself, computer, barcode reader, and screen. The equipment used for the evaluations presented here was a 'Image S', serial number 28 (05/2015) (weight 43 kg) (Figure 1).

The scanner consists by a helical vibratory system which orders the dehulled grains ensuring that each grain would



Figure 1. General view of the Image Rice Grain Scanner (without computer, screen and bar code reader).

Image - Rice Grain Scanner: a three-dimensional fully automated assessment of grain size and quality traits

pass individually throughout a special sensor and camera images are taken from each grain, allowing to get a complete coverage of the whole grain surface, which provides a more complete analysis. This is only possible because the images are captured during grains free fall. All pictures, from all grains are analyzed by the scanner and its software, but the system is set up to save only 100 pictures of each sample. This number can be changed by demand.

## OPERATION, TECHNICAL DETAILS AND NEW APPROACHES

### Sample preparation, identification & scanning

Rice samples can be identified by a field tag with a bare code, according field plots at the field trials. Epagri uses a model/system which was courtesy of Embrapa/CNPAF (Pinheiro et al. 2011), however with some changes. To prepare the samples for the analysis at the Image, each sample of rough (paddy) rice must be dehulled to produce brown rice, which is milled to produce white milled rice. At this point, the advantage of the software designed by Epagri/Selgron over the traditional evaluations of the milled yield quality will start to become evident.

Typically, as described by McClung (2003), 125g rough rice grains are used to determine milling yield (the separation into whole milled and broken grains). In traditional rice quality laboratories, the milling quality is estimated using a trieur that separate whole and broken grains, so that the they are weight separately, and the milling quality is given by the percentage of these weight related to the initial weight of the rough rice sample.

At Epagri's Breeding Program, 100g rough rice is used for starting milling quality analysis. It is obvious that, if a lower amount of grains is used, like 50 g or less, measurements of milling yield could be done a little bit faster, saving time immediately, and in a long term, saving time if this analysis could be done faster in earlier breeding generations, like F3. So an amount of 50g or 25g of rough rice can be used for milling the desired sample, to win the quality data at the Image Rice Grains Scanner starting from a final sample of, respectively, about 35g and 21g milled rice. Of course, this smaller amount of sample would save time. It is important to stress out that the Image is set up to an initial sample of rough rice, of 100g, 50g or 25g. No other amount can be used, since these are the weights of rough rice considered by the software to estimate the milling quality (milling yield), that means, percentage of whole and broken grains based on the normal rough rice starting sample. Image do not weight the samples, but instead use a medium kernel density to estimate the weight, and so, to estimate, later on, the amount of whole and broken kernels as a milling quality trait. If there are no interest in milling quality data, of course each other initial amount of sample can be used for the remaining outputted parameters. After dehulled, such amount of sample should easily pass to a kitchen sieve to eliminate small particles or dust before going into Image's helical plate.

The sample is finally placed at Image's helical plate at once (Figure 2). The scanner is started pressing a button at the screen and in a few minutes all the grains of the whole sample are analysed. The engine orders the dehulled grains ensuring that each would pass a sensor individually,



Figure 2. Image running a milled rice sample at the helical vibratory plate.

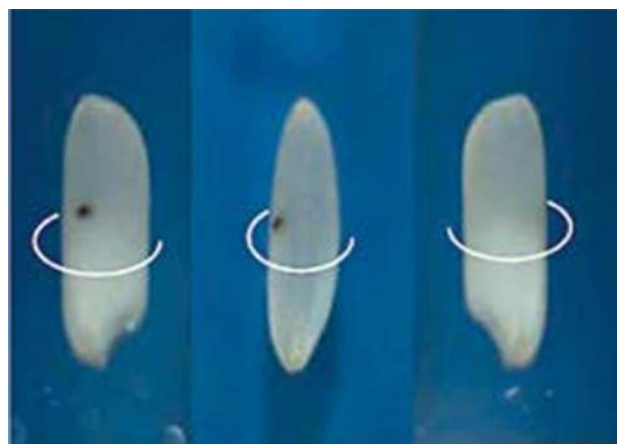


Figure 3. Three-dimensional images during the free fall allowing an analysis that covers the whole grain surface.

in free fall, while the grain is shot from different sides by a camera, so covering the whole surface of each grain (three-dimensionally), which seems to be a unique feature for rice scanners at the market (Figure 3). The software allow a precise measurement of grain geometry, shape, size, grain classes, chalkiness, other defects of the grain, and an estimation of milling quality, resulting in the output of 39 different information.

After finishing the grain analysis, the sample tag is read by the bar code reader to record the exact sample number which is automatically insert into the table (spreadsheet) that is seen at the computers screen. If necessary samples number or name could be inserted or edited manually.

A sample run at the Industrial version of the Image scanner (Selgron) can be found at YouTube looking for ‘SELGRON - Classificador de Graos – Analyzer’.

### Running the software platform

All the titles and buttons at the software interface are still in Portuguese but will be available in English by demand. Once finished the analysis (run) of a sample at the scanner, the data will appear at the spreadsheet on the screen at a clear six kind of columns division. The first one is composed by columns with traits given according the trial automation system (adapted from Pinheiro et al. 2011). This can be filled up with field data later one. The second division is related to sample identification (Black header ‘INFORMAÇÕES’, Figure 4), like complete code of the sample, followed by the trial code, and the plot (sample) number itself. The third division of data output (green header Sample Data, ‘AMOSTRAL’, Figure 4) gives information about the size of the whole sample (length, width and thickness), including all grains (even broken ones), standard deviations (of length, width and thickness), length/width relation, and total number of grains analysed. Also the fifth and sixth division of columns will be filled up with data, that means, respectively, the brown header (‘TOTALIZADOR’), which contain the different classes (according Brazilian norms, but can be changed by demand according to needs or country norms) of the analysed grains, and the last black header, with the columns of the milled quality data (‘RENDIMENTO DE ENGENHO’) (Figure 5).

At this point, is even possible to apply a ‘filter’ to the already obtained data just before saving each sample. If a filter, that means, a limit value for length is inserted, and a command is chosen ( $\geq$  or  $>$ ,  $\leq$  or  $<$ ), this will fill up the fourth group

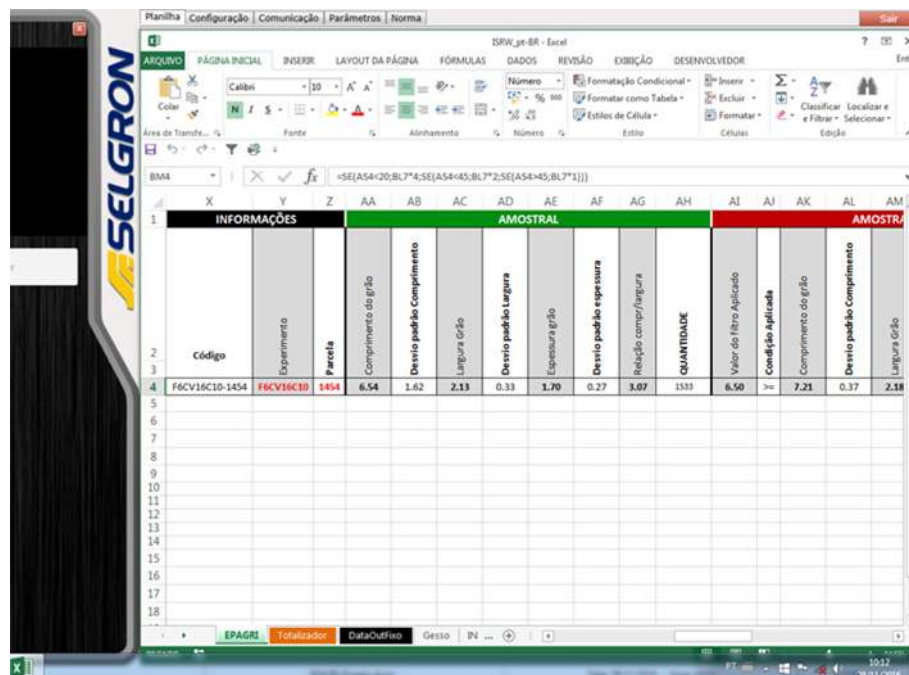


Figure 4. Partial screen view showing the samples information header columns (black), ‘amostral’ header columns (green) with the grain size data, and the initial part of the ‘amostral com filtro’ header columns (red) showing grain size data after a use of a limiting filter.

Image - Rice Grain Scanner: a three-dimensional fully automated assessment of grain size and quality traits

of columns (red header: Sample with filter) of the output spreadsheet just in time. The filter is useful to determine the real size of the unbroken grains, that means, to know the grain length (mm), width (mm) and thickness (mm) without considering the measures of the broken grains. If a filter of 6.5 mm is applied to the length and the command  $\geq$  is chosen (Figure 6), it means that the software will recalculate the average of the three grain dimensions (length, width and thickness) taking in account only the grains over 6.5 mm of length. This will give the breeder a real idea of new line/cultivar grain's

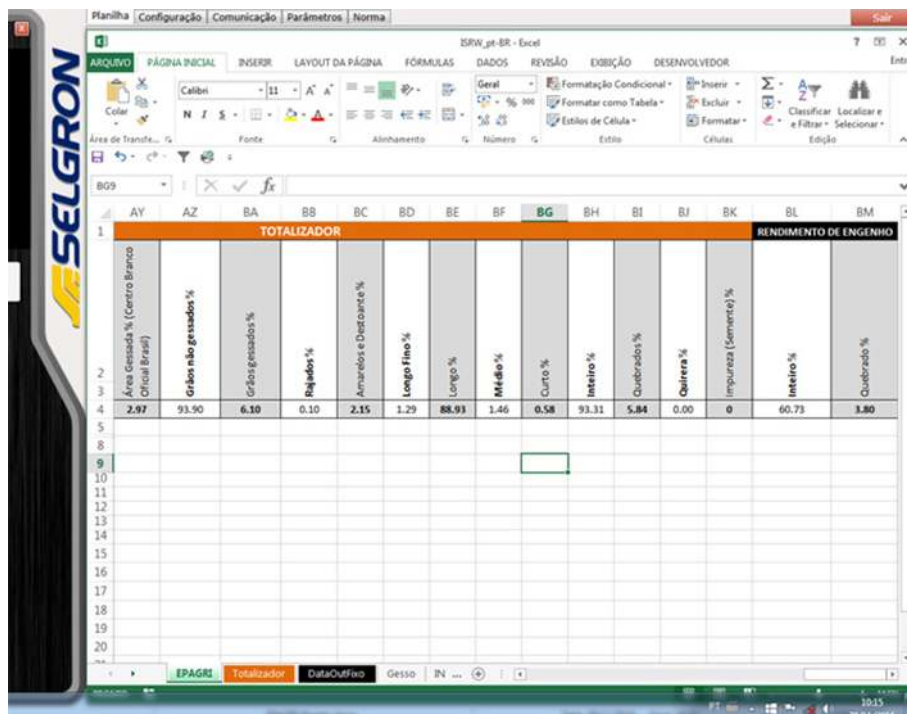


Figure 5. Partial screen view showing the TOTALIZADOR columns (brown header), with different classes of the grains, and the milling quality columns (last black header).

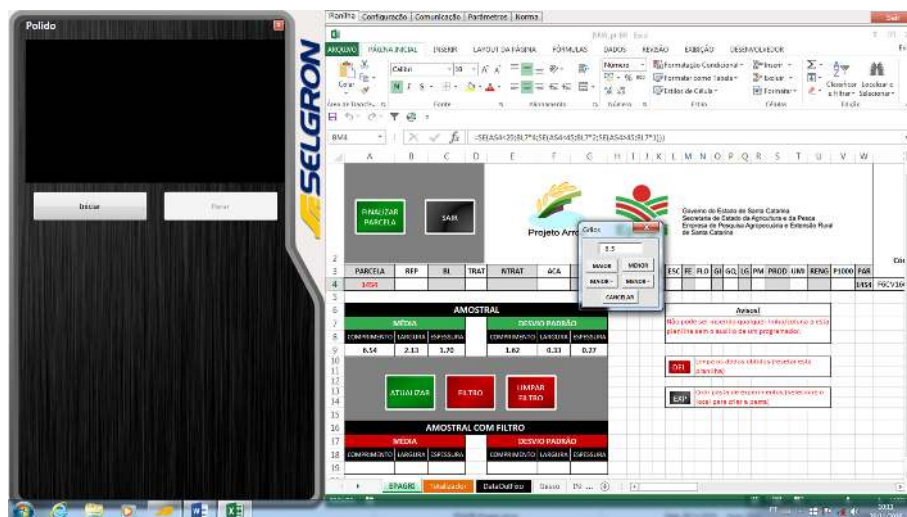
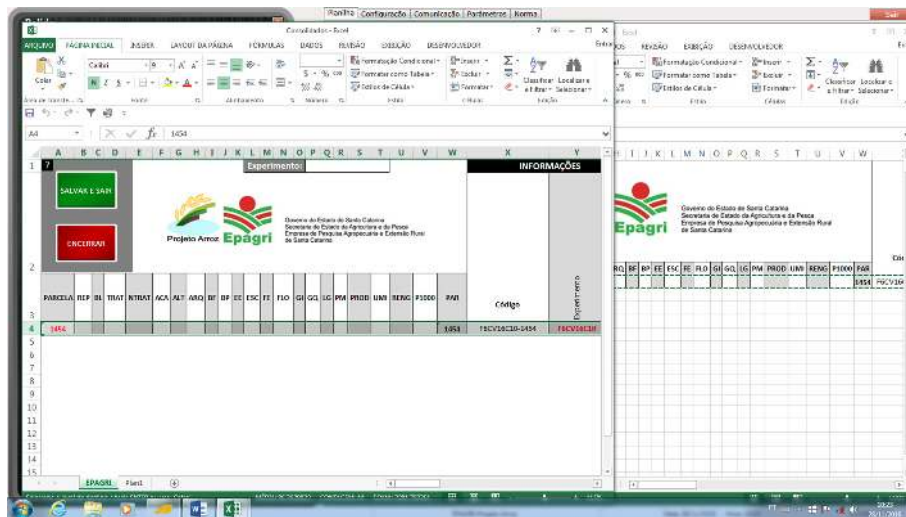


Figure 6. The whole screen view showing the average and standard deviation of the grain size of a sample (green headers), and the small screen where the value of a filter for the length is introduced by the operator. It will be showed again in the line of each respective sample.



**Figure 7.** Screen view after the first sample was analysed, showing a green button, to be save the data and proceed to analyse a second sample, or the red button, if all samples of a trial were analysed and a consolidate file with all data should be generated.

size. Broken grains can not always be considered as a defect from a genotype, but sometimes is a result of other factors. The use of the filter eliminates it from the grain size estimation, because broken ones are usually extracted during rice industrialization. Of course, just as it is known that the number of broken grains is not always the fault of the genotype itself, it is also known, in other turn, that the milling quality could be different for different genotypes. To know it, as a trait, is very important for breeders, but on the other hand, milling quality is also affected by environment conditions, that means, could be different in different lines under different situations, but this is handled by the Image by milling quality estimation (based on whole and broken grains). Instead and beside this, breeders are normally interested in the real size of unbroken grains, and this can be measured by the Image excluding, by applying this filter, the broken grains, since there are broken not necessary because of its genetic background.

After that it is necessary to save the first sample of a trial by indicating in which folder this and all other next samples from the same experiment should be saved. 'Plot finalize' green button (Figure 6) must be pressed (FINALIZAR PARCELA). A new window will open (Figure 7) showing the first analysed sample, and two buttons will be there as follow: a green one (SALVAR E SAIR), to save the data and go out to proceed a second sample, which in this case could be placed on the helical plate starting a new analysis. Or a red button, to 'end' (ENCERRAR) this experiment/trial. This button must be pressed right after the analysis and saving of the last trial. This will group all analysed samples and its data into the same Excel spreadsheet called 'consolidate file'. Once pressed the red button 'Encerrar' (Figure 7), the consolidate date of all samples will be opened at a new window, and then the folder of the trial must be indicated again. Automatically all data will be saved as a consolidate Microsoft excel file, and a pdf file into the indicated folder.

## RESULTS AND DISCUSSION

The Image Scanner takes about 6 minutes to analyse 70g of milled rice (starting with a sample of 100g rough rice), and 3 minutes for 35g milled rice (starting with 50g rough rice). Since 35g milled rice usually represent about 1,000 grains, it is a good representative sample of a line or variety. According to Epagri's experience over the last growing season, a sample of 50g rough rice is a good deal between accuracy and analysis spent time.

To clarify, next can be find the 39 information given by the Image (Scanner outputs) at the consolidate spreadsheet. Sample identification columns (black header): 'Complete Field trial code'; Field experiment/trial code; Plot (Sample) number (inside each trial). Sample grain size (green header): grain length (mm), standard deviation of grain length, grain width (mm), standard deviation of grain width, grain Thickness (mm), standard deviation of grain thickness, length/width Ratio, number of grains analysed.

Applied settings for the filter will fill up the next group of columns, called 'Amostrado com filtro' (Sample with filter) (red header) composed by: grain length (mm), standard deviation of grain length, grain width (mm), standard deviation of grain width, grain thickness (mm), standard deviation of grain thickness, length/width Ratio, number of grains analysed.

The 'Totalizador' (totalizer) group of columns refers to the different classes and/or defects according the Brazilian Ministry of Agriculture, Livestock and Supply (IR 06 MAPA). Adjustments at the class building parameters can easily be done by Selgrom regarding to the needs or specific regulations of each country (Figure 8). The totalizer columns are composed by: estimated sample weight, good grains (%), white bally grains (%), moldy and burnt grains (%), chopped/stug and stained grains (by insects or fungi) (%), total chalked area (%), no chalked grains (%), chalked grains (%), striped, yellow and color outsiders (%), long-thin grains (%), medium grains (%), short grains (%), whole grains (%), broken grains (%), very small grain pieces (%), impurities (%).

The last group of columns at the output spreadsheet file are the Milling Quality columns (black header), given by whole grains and broken ones percentage. It is very important not to confuse the apparently twice appearance of whole and broken percentage grain at the output spreadsheet. At the 'Totalizer' columns the percentage of whole and broken grains are the really exact division of whole and broken grains of the sample based on number of whole and broken grains. On the other hand, whole and broken grains at the Milling Quality Columns are an estimate of milling quality based on weight of whole and broken grains, related to an initial sample of 100, 50 or 15 g rough rice. In this case are even included has broken ones the short grains, which would be anyway treated by the industry as broken ones, and discarded, like it would be passing a trieuer in conventional milling quality analysis.

With reference to the milling quality, it will be discussed data of an, at Brazil, official experiment called 'Value for Cultivation and Use (VCU)' carried out at the Epagri - Itajaí Experiment Station (2015/2016 grow season), in which 20 lines and 3 rice cultivars were evaluated under a random block design with three replications. If final milling quality estimated by the Image scanner is compared to the milling quality ('Renda') given by the traditional method {(% of milled rice grains x (% of hulls + bran stream))}, a correlation of 88% was found. At average, the total milling quality of the conventional method (using trieuer and weight) estimates at 0.85% more milled rice (whole and broken) if compared to the amount estimated by the scanner. Looking at the conventional average milling yield the 69 observations result at 70.67%, while the average milling yield coming from the Image analysis was 69.82%. Otherwise, if the 69 samples are compared for milling quality in the two methods the average between the differences of the pair of observations (conventional and image), for whole grains, was -1,31%, which means that the scanner data for whole grains result always

Classe	Altura maior ou igual que	Altura menor que	Relação altura/largura maior ou igual que	Espessura menor ou igual que
Longo Fino	6.0	=	2.8	1.9
Longo	6.0	=	Relação altura/largura menor que	Espessura maior que
Médio	5.0	6.0		
Yemetha	4.5	=	2.5	2.1
Curto	4.5	5.0		
Quebrado	2.1	4.5		
Querera	0.0	2.1		
Canjica	3.1	=		
Canjica	2.1	3.1		
	5			

Figure 8. Parameters for the build of the different classes at the 'Totalizer' group of columns at the output spreadsheet that can be adjusted according needs or country norms.





Image - Rice Grain Scanner: a three-dimensional fully automated assessment of grain size and quality traits

could be done at very early breeding stages. Probably this equipment could also be successfully used in phytopathology or entomology studies, or even could be applied to studies under abiotic stresses like cold or heat at reproductive stages, since many defects can be measured, and so, a decrease in quality traits can be accurately detected and measured.

## ACKNOWLEDGEMENTS

This technology was supported by the 'PAC Embrapa' (Convênio 732209/2010 – 10.200.10/0104-5 - Modernização da Infraestrutura Física das Organizações Estaduais de Pesquisa Agropecuária). Thanks to 'Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina' (Fapescc) for supporting the automation of rice field trials at Epagri. The authors are grateful to Mr. Gildo João Sardo (Epagri) for support. Thanks to Dr. Adriano P. de Castro and Sérgio Lopes Júnior (Embrapa/CNPAP) for the kind support at initial stages of automation of rice field trials efforts at Epagri-Itajaí Experiment Station.

## REFERENCES

- McClung AM (2003) Technique for development of new cultivars. In Smith CW and Dilday RH (eds) **Rice, origin, history, technology, and production**. John Wiley & Sons, Hoboken, p. 177-202.
- Pinheiro PV, Junqueira RBM, Junior SL and Pedrochi MA (2011) **Processo gestão de experimentos: a experiência da Embrapa Arroz e Feijão**. Embrapa, Santo Antônio de Goiás, 36p. (Documentos 266).
- Tanabata T, Shibaya T, Hori K, Ebana K and Yano M (2012) SmartGrain: High-throughput phenotyping software for measuring seed shape through image analysis. **Plant Physiology** **160**: 1871-1880.
- Whan AP, Smith AB, Cavanagh CR, Ral JF, Shaw LM, Howitt CA and Bischof L (2014) GrainScan: A low cost, fast method for grain size and colour measurements. **Plant Methods** **10**: 1-10.