scientific reports

Supplementary

How Germination Changes During Individual Seed RGB-Space Differentiation: The Case of *Pinus sylvestris* L. cv. Negorelskaya

Tatyana P. Novikova [1], Paweł Tylek[2,*], and Arthur I. Novikov [3*]

[1] Faculty of Computer Science and Technology, Voronezh State University of Forestry and Technologies named after G.F. Morozov, 8 Timiryazeva, 394087 Voronezh, Russia; novikova_tp.vglta@mail.ru

[2] Faculty of Forestry, University of Agriculture in Krakow, Al. 29 Listopada 46, 31-425 Kraków, Poland; rltylek@cyf-kr.edu.pl

[3] Department of Agricultural Technologies Management and Agromonitoring, Agrophysical Research Institute, 14 Grazhdanskiy av., 195220 Saint Petersburg, Russia; nvatdo@gmail.com

* Correspondence: rltylek@cyf-kr.edu.pl (P.T.); nvatdo@gmail.com (A.I.N.)

Appendix A

The library of forest reproductive material (FRM-Library) [73], which is based on Pravdin's conjecture [11,24], is expanding and being filled with data on individual parameters and indicators of each single seed of *P. sylvestris* L. cv. Negorelskaya.

The relational data model (enlarged) of the individual "seed–culture" technological passport (*P. sylvestris* L. cv. Negorelskaya) are presented in Figure A.1.

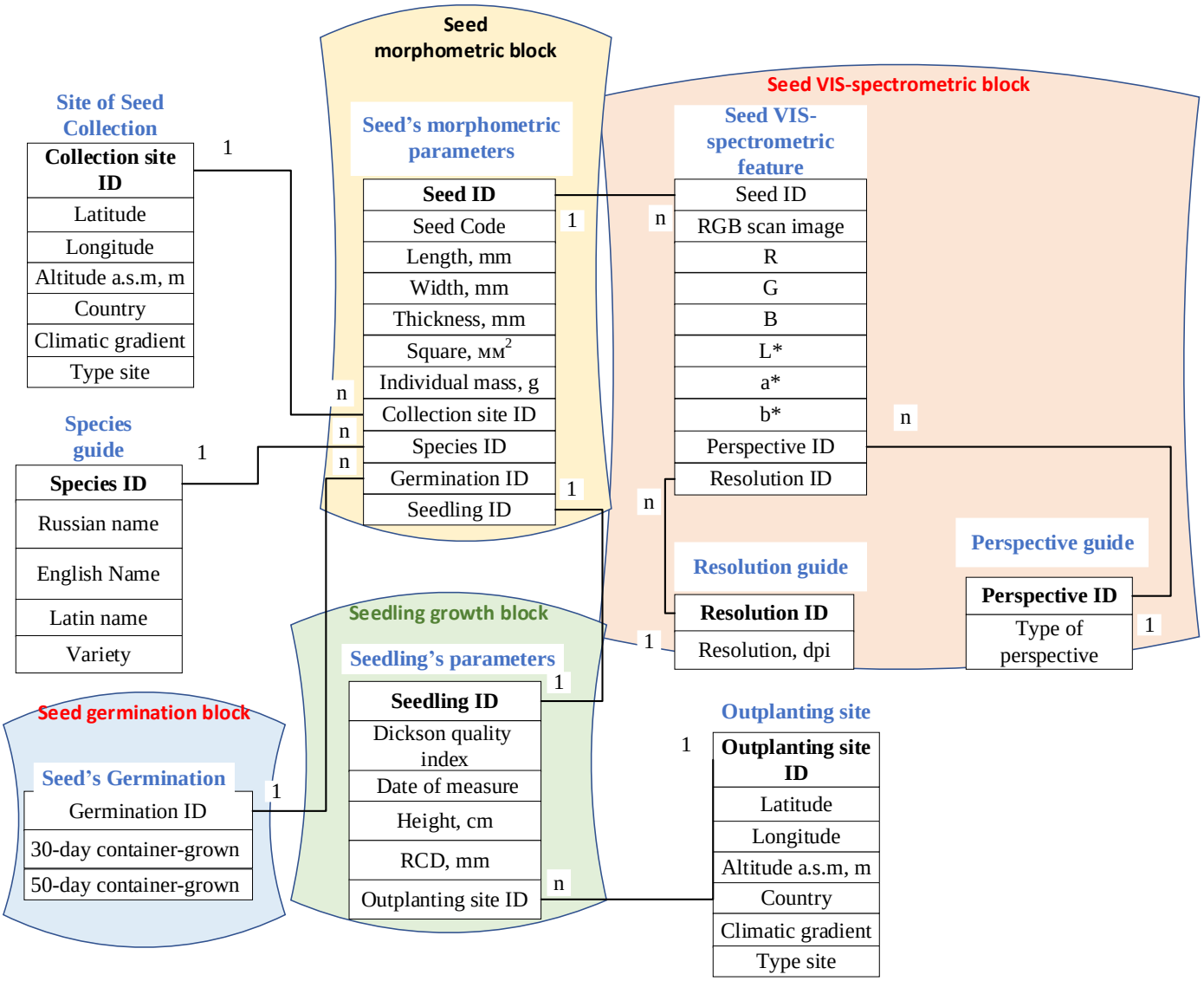


Figure A1. The relational data model (enlarged) of the individual "seed–culture" technological passport (*P. sylvestris* L. cv. Negorelskaya). This relational model takes into account the morphometric properties, including individual seed mass (seed morphometric block) and VIS spectrometric properties, of a single seed in different VIS color channels (seed VIS spectrometric block).

Dataset blocks used in this study (see Figure A1) are include:

- Seed morphometric block (from Dataset 1: Morphometric data of individual seeds (N = 1200)) of the "Negorelskaya" *Pinus sylvestris* L. (empirical dataset). The tabular dataset represents the results of direct measurements of the geometric dimensions (length, width and thickness in mm) and mass in grams of each individual seed, as well as the calculated values of the projection area and volume of the described ellipsoid on the basis of these parameters. The dataset allows for correlation and regression analyses between geometric parameters and individual seed weight and can also be linked to other datasets in the FLR-Library [24] to form summary queries.

- seed VIS-spectrometric block (from Dataset 2: VIS-Spectrometric data of individual seeds (N = 1200) of the "Negorelskaya" *Pinus sylvestris* L. (empirical dataset)). The file dataset represents the results of direct scanning by a charge-coupled device of 30 seed groups (the number of seeds in group n = 40 and their location on the scanner glass coincides with their future location in a side-slit container during sowing) in the RGB color space of the visible (VIS) spectrum with resolutions of 300, 600 and 1 200 dpi. The dataset can allow scientists to use the [Particle Analysis] add-in of the FiJi open-source software to segment the image of an individual seed (for example, as in Bernardes et al. (2022) [59] or A. Loddo et al. (2023) [74]) and obtain quantitative data of two projections of a single seed in integrated RGB spaces or separate channels of the visible region of the spectrum. The dataset can also be linked to other datasets in the FRM-Library [24] to form summary queries and study the effects of VIS-spectrometric seed parameters on germination and early growth;
- Seed germination block (from Dataset 3: Germination data of individual seeds (N = 1200) of the "Negorelskaya" *Pinus sylvestris* L. (empirical dataset)). The tabular dataset includes the results of container-grown germination studies of each of the 1200 seeds on the 30th and 50th days in 120 cm^3 cells of 40-cell side-slit containers filled with peat substrate and mulched with perlite. The dataset allows for correlation and regression analyses of the effect of individual seed parameters on the indicator of seed sowing qualities and can also be linked to other datasets in the FLR Library to generate summary queries to predict the effect.

A.1. Seed morphometric block

The seed morphometric block (from Dataset 1) contains tabular data of 1 200 single seeds of the *Pinus sylvestris* L. cv. Negorelskaya. The data on the block are generated in accordance with Table A1.

Table A1. The table data format of the *23-26-00228-RSF-DataSet-SeedNG-Morphometry.xlsx* file from Dataset 1. Total number – ordinal global seed number (from 1 to 1200), key field, data type – numeric, variable type – categorical, nominal; Sample's number – number of a sample (batch) of seeds selected from different parts of the seed pile by quartering [42], data type – numeric; variable type – categorical, nominal (takes values from 1 to 3); Seed mass: data type – numeric; variable type – quantitative (direct measurement using analytical scales); Seed length: data type – numeric; variable type – quantitative (direct measurement using a micrometer); Seed width: data type – numeric; variable type – quantitative (direct measurement using a micrometer); Seed thickness: data type – numeric; variable type – quantitative (direct measurement using a micrometer); Seed square: data type – numeric; variable type – quantitative (calculation according to subsection 4.3); Ellipsoid volume: data type – numeric; variable type – quantitative (calculation according to subsection 4.3).

Total number	Sample's number	Seed mass, mg	Seed length, mm	Seed width, mm	Seed thickness, mm	Seed square, mm^2	Ellipsoid volume, mm^3
1	1	0,013	4,50	2,57	1,61	36,31	9,74
…	…	…	…	…	…	…	…
1166	3	0,0102	5,25	2,75	1,50	45,33	11,33
1167	3	0,0055	4,02	2,44	1,19	30,80	6,11
1168	3	0,0065	4,50	2,51	1,17	35,47	6,92
…	3		…	…	…	…	…

A.2. Seed VIS-spectrometric block

Seed VIS spectrometric block formed from Dataset 2: "VIS spectrometric data of individual seeds (N = 1200) of the "Negorelskaya" *Pinus sylvestris* L. (empirical dataset)". The block contains file data of 1,200 individual seeds, which are arranged in scans of 40 seeds according to the number of cells in the container. The formation of file data in the block occurs as follows.

Three main folders of the file dataset were formed in accordance with Figure A.2, corresponding to three samples of *P. sylvestris* L. cv. Negorelskaya seeds selected for the study according to subsection 4.1 of main paper.

- 1NG(1-400);
- 2NG(401-800);
- 3NG(801--1200).

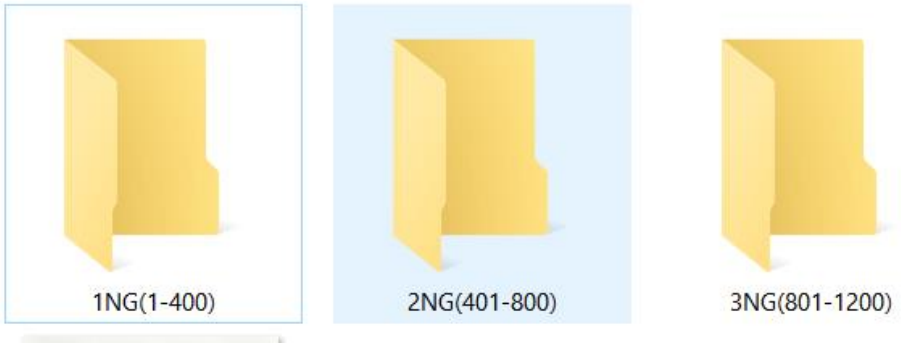


Figure A.2. The structure of the first level of the file dataset containing VIS scans of *P. sylvestris* L. cv. Negorelskaya seeds for the study of spectrometric parameters.

Inside each folder in Figure A.2, three second-level folders were formed in accordance with Figure A.3:

– XNG@300;

– XNG@600;

– XNG@1200; here, X is the sample number (1, 2 or 3).

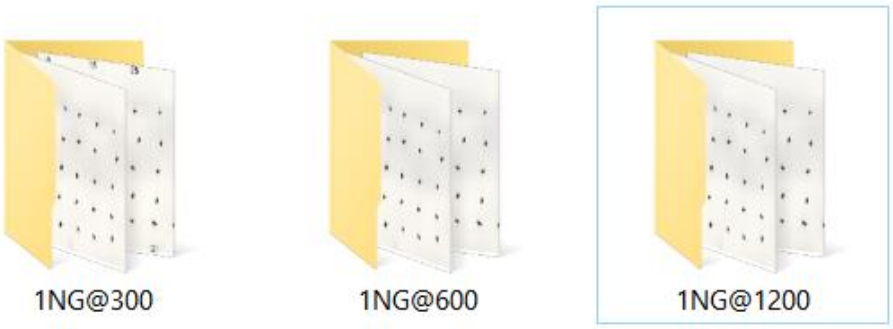


Figure A.3. The structure of the second level of the file dataset using the example of the folder 1NG(1--400) containing VIS scans of Scots pine (*P. sylvestris* L. cv. Negorelskaya) seeds for the study of spectrometric parameters: @300, with the resolution of obtaining an optical image in the visible wavelength range of 300 points/inch; @600–600 points/inch; @1200–1200 points/inch.

Inside each folder of the second level, XNG@YYY, the file data were formed according to Figure A.3 in accordance with Figure A.4:

– dZ(NS-NF)@YYY=Scan;

– vZ(NS-NF)@YYY=Scan.

Here, d is the dorsal side of the seeds; v is the ventral side (with a 180 degree rotation) of seeds; Z is the side-drain number of the container for subsequent seed sowing; NS is the initial seed number in the container; NF is the final seed number in the side–drain container; YYY is the resolution of the VIS scan obtained in the optical wavelength range (300, 600 or 1200).

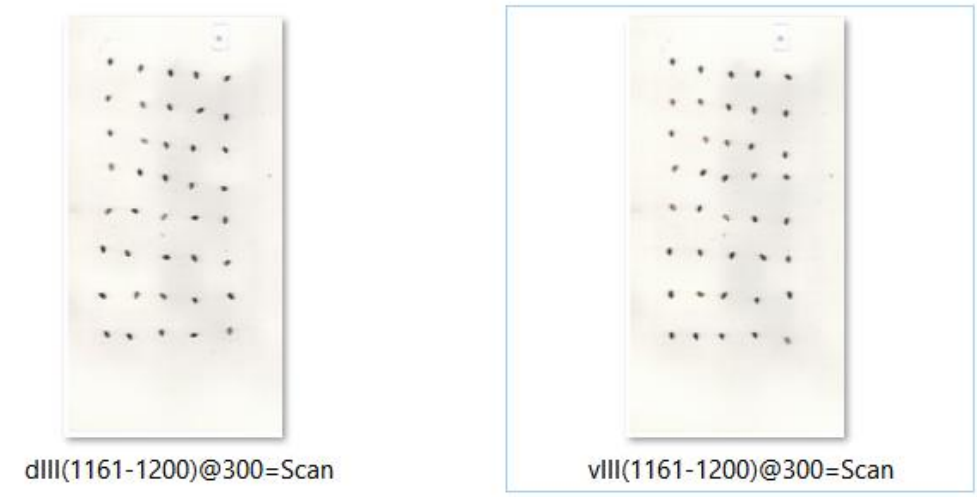


Figure A.4. The structure of the third level of the file dataset uses the example of the 3NG@300 folder, which contains VIS scans of Scots pine seeds (*P. sylvestris* L. cv. Negorelskaya) for the study of spectrometric parameters: the Roman numeral indicates the container number.

A.3. Seed germination block

Seed germination block formed from Dataset 3: "Germination data of individual seeds (N = 1200) of the "Negorelskaya" *Pinus sylvestris* L. (empirical dataset). The block contains tabular data container-grown 30-day germination of each of the 1 200 individual seeds. The data in the block are generated in accordance with Table A.2.

Table A.2. The table data format of the *23-26-00228-RCF-DataSet-SeedNG-G30.xlsx* file from Dataset 3. Total Number – ordinal global seed number (from 1 to 1200), key field, data type – numeric, variable type - categorical, nominal; Sample's number – number of a sample (batch) of seeds selected from different parts of the seedlot by quartering, data type – numeric; variable type – categorical, nominal (takes values from 1 to 3); Seed's number – the number of the seed inside the sample (batch) selected from the seed pile, data type – numeric; variable type – categorical, nominal (takes values from 1 to 400); Container's number – the number of the container in which the studied seeds were sown, data type – numeric; variable type – categorical, nominal (takes values from 1 to 40); G30 – germination index of each individual seed on the 30th day from the moment of sowing in a side-slit container, data type – numeric; variable type – categorical, rank (takes only two values 0 – there is no seed germination; and 1 – there is a viable plantlet).

Total Number	Sample's number	Seed's number	Contayner's number	G30, (0-No | 1-Yes)
…	…	…	…	…
480	2	80	12	0
481	2	81	13	1
482	2	82	13	1
483	2	83	13	1
484	2	84	13	1
485	2	85	13	1
486	2	86	13	0
487	2	87	13	0
488	2	88	13	1
489	2	89	13	0
490	2	90	13	1
491	2	91	13	1

Total Number	Sample's number	Seed's number	Contayner's number	G30, (0-No | 1-Yes)
492	2	92	13	1
493	2	93	13	1
494	2	94	13	1
495	2	95	13	1
496	2	96	13	0
497	2	97	13	0
498	2	98	13	1
499	2	99	13	1
500	2	100	13	1
…	…	…	…	…

* The gray color of the cell represents the beginning of another container, ** the red color of the cell clearly demonstrates the absence of a seedling (zero germination of the seed).

A.4. Seed morphometry processing

For each seed from three sets (total number of seeds N = 1200), the dimensions, weight, area, and volume of the ellipsoid were measured according to the methodology developed on the basis of [59,74,75] and placed in transparent pockets under an individual number.

The individual weights of the seeds were recorded via special laboratory analytical scales. The average temperature and humidity in the laboratory during the study were 25 °C and 21%, respectively. The weight of each seed was recorded via laboratory analytical scales with an accuracy of 0.0001 g. Before measurement, the scales were installed with the possibility of excluding the effects of vibration, heat sources, air flows and sudden temperature fluctuations, which were balanced via an integrated bubble indicator, which was set to zero. Next, the measured seed was placed with tweezers in the center of the circle, transparent flaps were closed to prevent the influence of air movement, and the seed mass readings were recorded after stabilization of the corresponding arrow. The readings were recorded in a special journal.

The research methodology used for the determination of the geometric characteristics of the seeds of *P. sylvestris* L. cv. Negorelskaya. The parameters, such as the surface area of the seed (mm^2) and the seed volume (mm^3), were selected.

The surface area of the seed was calculated via the formula for the area of the ellipse, which most fully resembles the shape of the seed:

$$S_c = \pi \cdot L \cdot W, \quad (1)$$

where S_c is the surface seed square, mm^2; π is a constant equal to 3.14; L is the length of the seed, mm; and W is the width of the seed, mm.

The single-seed volume was calculated via the formula of the volume of an ellipsoid, which most closely resembled the shape and volume of the seed:

$$V_c = 4/3 \cdot \pi \cdot 0{,}5\ L \cdot 0{,}5\ W \cdot 0{,}5T, \quad (2)$$

where V_c is the volume of the seed, mm3; π is a constant equal to 3.14; L is the length of the seed, mm; W is the width of the seed, mm; and T is the thickness of the seed, mm.

A.5. Seed image processing

Scanning, according to the proposed method of the author T.P. Novikova [68] (Patent application RU 2024137297, 2024-12-12), was performed for 40 seeds, placing them on flatbed scanner glass with a white background [16] in the order of future sowing in containers. The seeds were removed from individual pockets with an individual number and placed on the scanner glass in accordance with Figure 25.

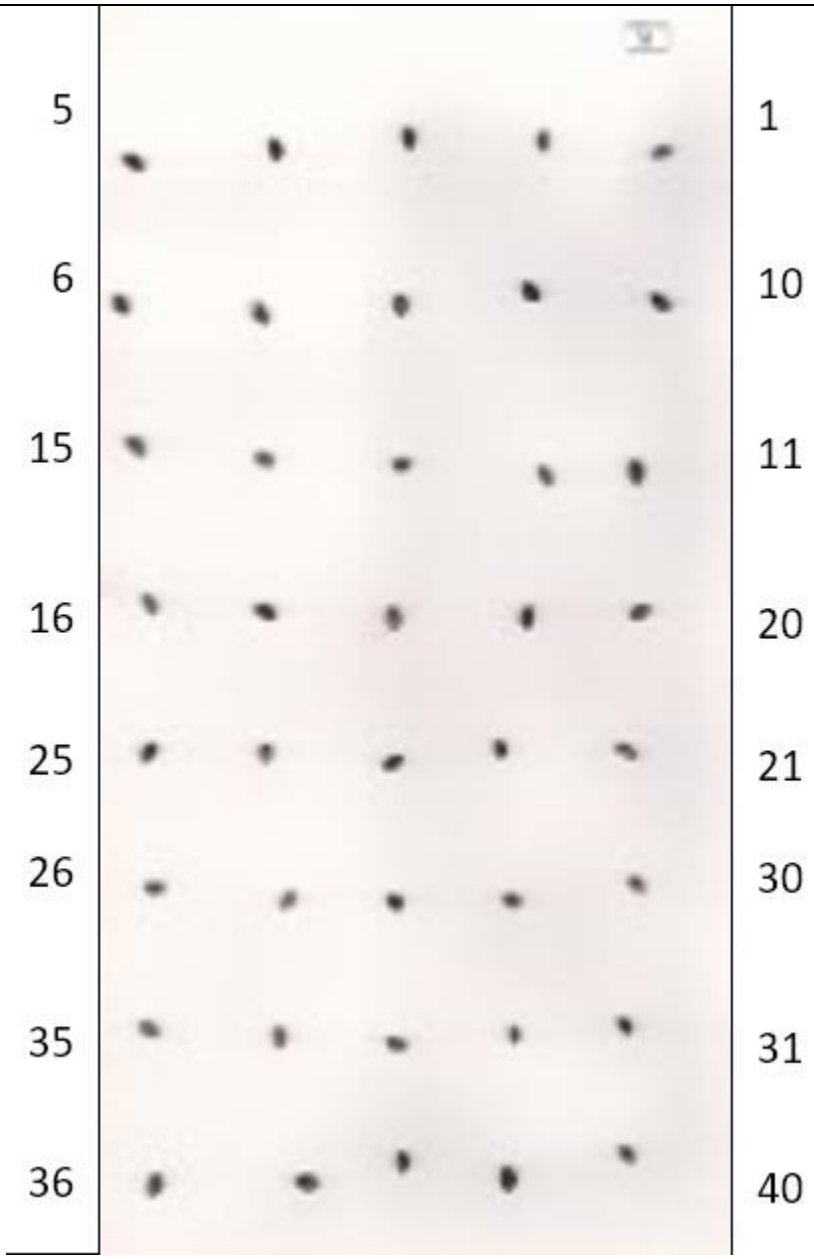


Figure A.5. A fragment of the location of the seeds on the flat-bed scanner glass. The **St**-mark was set for the beginning of the countdown. © Adapted from coauthor T.P. Novikova (2024) [68]

We preconfigured the field size of the 40-seed scan by clicking the [Preview] button in the scan window of the scanner interface (Brother DCP). The scan paper size was cut to 280*145 mm. For the subsequent study of spectrometric properties and to ensure a sufficient level of subsequent segmentation of the image of the dorsal and ventral projections of the seeds of Scots pine (*P. sylvestris* L. cv. Negorelskaya) with a square of 124 × 124 pixels, for example, as in Rodrigo K. Bernardes and coauthors [59], a sufficient level of randomization, as well as minimization of the noise of the CCD scanner matrix, provided for the location of the seeds of Scots pine (*P. sylvestris* L. cv. Negorelskaya) at a distance of at least 20 mm from the edge of the tablet in the order corresponding to the order of subsequent sowing of seeds in side-ingot containers.

The scanning resolution was set to 300 dpi, the scanning mode was color, and the brightness was set to the default value. Moreover, the paper size in the pixels was set to 1718*3309 pixels. Next, the [Scan] button was pressed, and the scan timing was determined via a smart stopwatch background, the value of which was entered into the Excel table. For the period of sample scanning, the time from the appearance of the "Data Transfer" window to the appearance of a thumbnail of the scanned seed image in the left menu

of the ABBYY Fine Reader program was assumed. The resulting seed scan was saved in uncompressed TIFF format with a file name of the form dI(1-40)@300=Scan, where d(v) is the conditional dorsal (ventral) orientation of the seed relative to the scanner glass; I (II, III) is the number of random samples of seeds from the seedlot; 1--40 is the unique seed cipher in the current study; @300 (600,1200) is the scanning resolution, with dots per inch; and =Scan is the color of the reflective substrate of the scanner or colored paper.

The resulting file (image) has the following numbering of each individual seed. After a sample of 40 seeds corresponding to the future location in side-ingot containers was scanned, the resolution and size of the paper were changed to 600 dots per inch and 3436*6619 pixels, respectively. Moreover, the timing was determined for a resolution of 600 dpi, and the data were entered into the corresponding cell of the Excel table. After the scan was saved, the resolution and paper size were changed to 1200 dpi and 6873*13238 pixels, respectively.

Thus, for each orientation of seeds with a certain color of the substrate intended for sowing in one container (1--40), three files with resolutions of 300, 600 and 1200 dpi were obtained (Figure A.6).

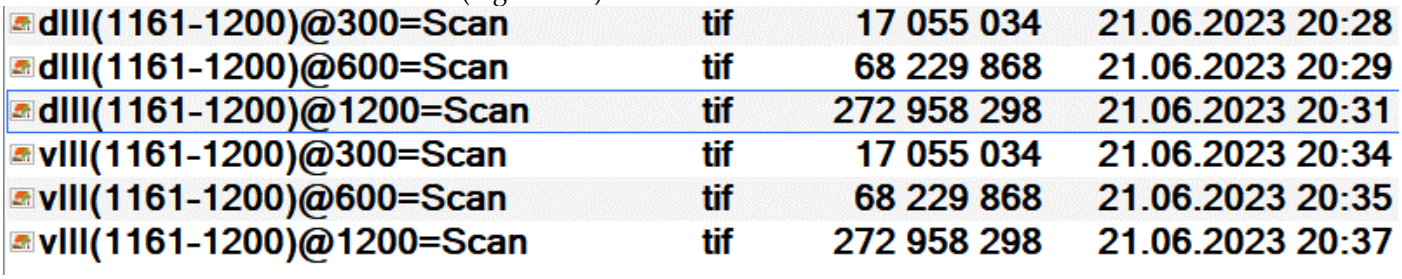
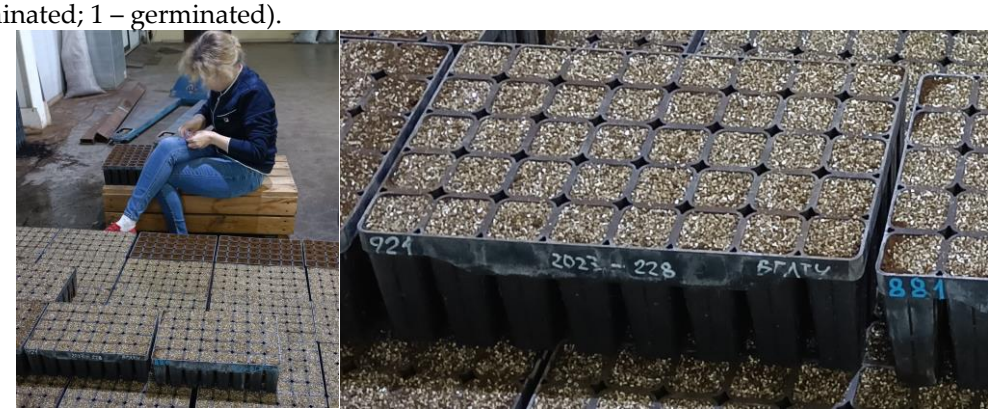


Figure A.6. A fragment of the listing of files (Total Commander) obtained by scanning seeds 1161--1200 intended for sowing in a 30-number container.

2.5. Seed germination processing

The seeds (three samples of four hundred seeds each) were sown manually on June 23, 2023, into each of the 40 cells with a volume of 120 cm^3 in HIKO V-120 SideSlit containers (size 352*216*110 mm, 526 seedlings per square meter; BCC AB, Sweden). Each container was prefilled with an acid reaction peat substrate, and the seed was placed in the center of the cell at a depth of 0.5–1 cm. The location of the seeds for subsequent identification was carried out in accordance with Figure A.7, *a* indicating the initial reference cell from the outside with a special marker as in Figure A.7, *b*. After sowing 40 seeds, each container was filled with mulch in the form of perlite and placed on a pallet for transportation to the greenhouse. Each sample of 400 seeds was placed in 10 containers. Thirty [42,43] days after seeding, the individual germination of each seed was calculated (0 – not germinated; 1 – germinated).



(a) **(b)**

Figure A.7. Germination analysis used in this study: general appearance (**a**); container labeling (**b**) © Photo by coauthor T.P. Novikova