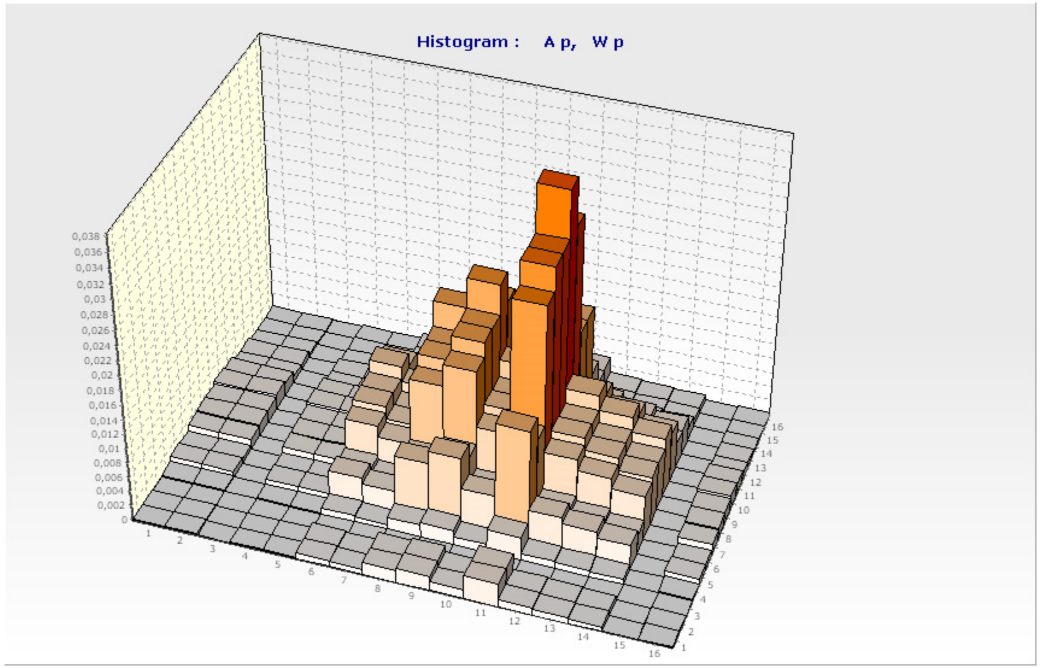
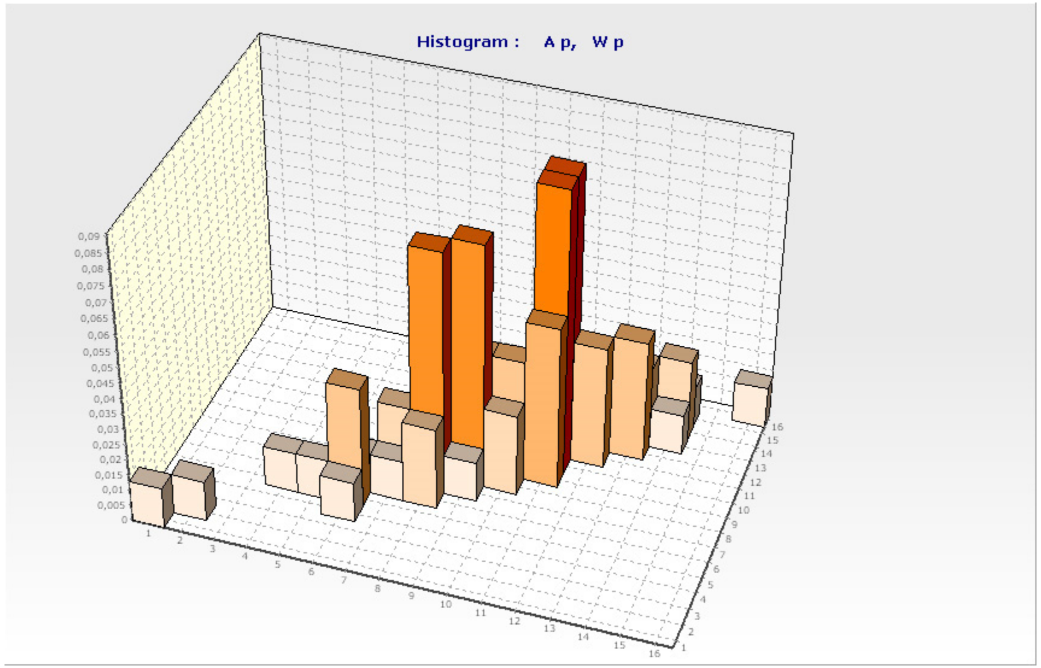
Statistically Dependent Random Variables

- Theoretical background
- Software HistAn2D and HistAn3D
- Examples

Statistically Dependent Input Variables

Some of input variables are **statistically dependent** however, e.g., cross-section characteristics, strength properties etc.





Statistics Analyze of the Measured Data

Pearson's correlation coefficient

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

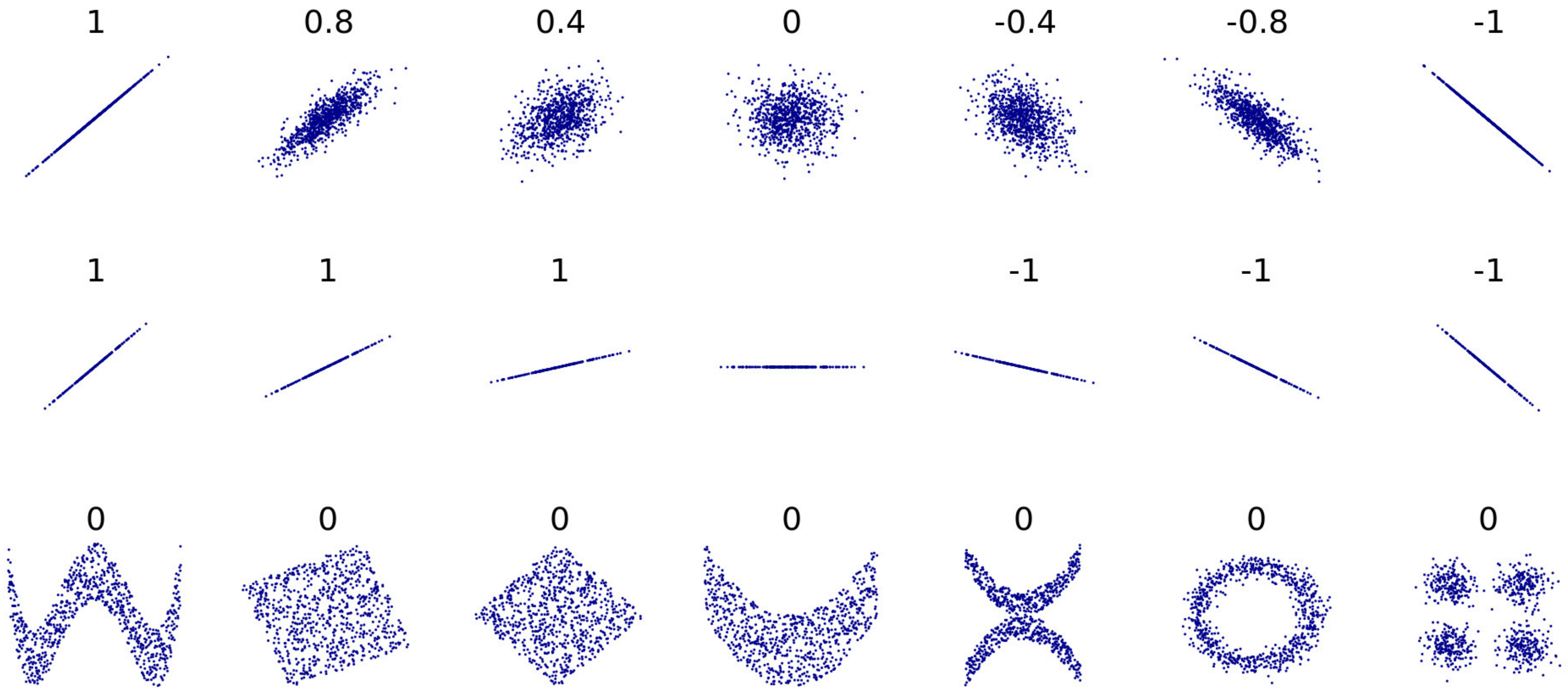
where x_i and y_i are elements of the both random quantities and $\bar{x}$ and $\bar{y}$ are mean values of those quantities

Spearman's coefficient of sequential correlation

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^{n}(p_i - q_i)^2}{n \cdot (n^2 - 1)}$$

can be determined by arranging n values x_i and y_i of the both random quantities by their size and by allocating sequence numbers p_i and q_i.

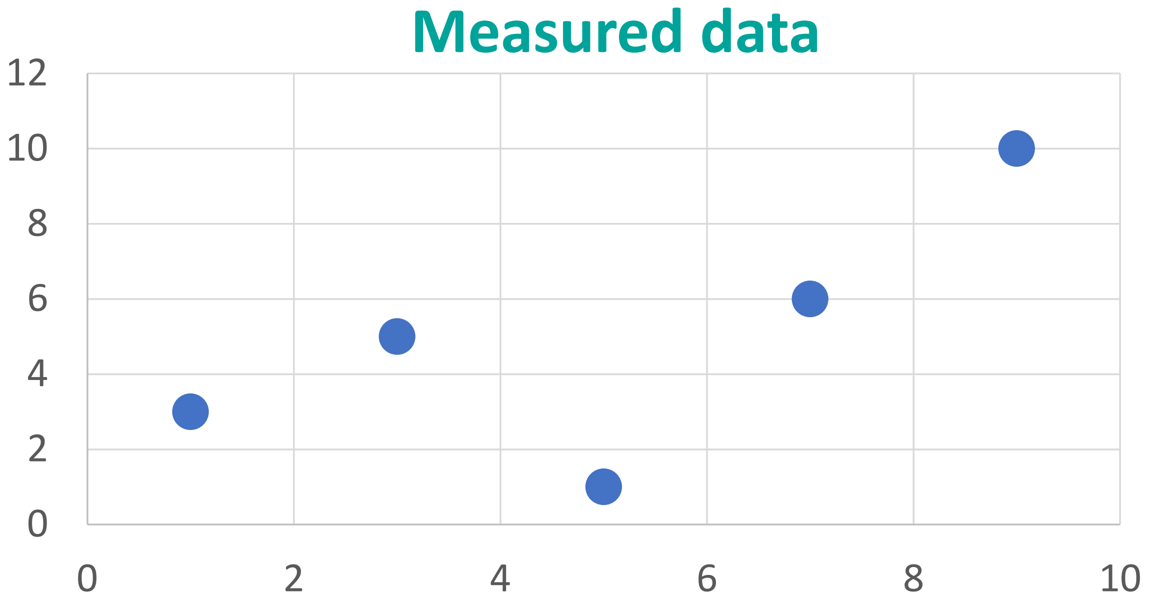
Statistics Analyze of the Measured Data



Example 1

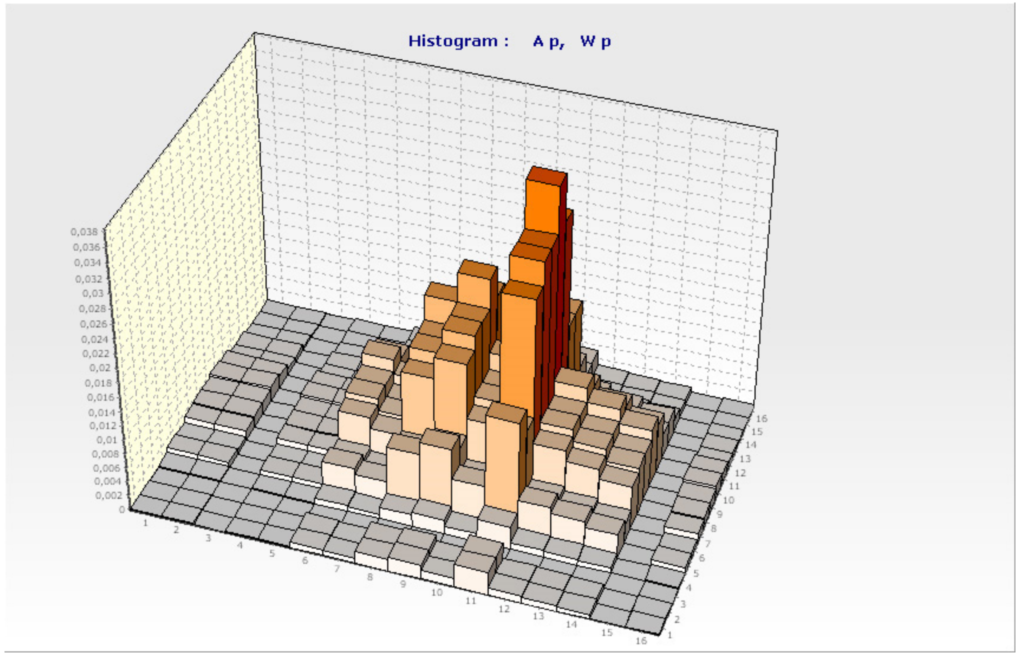
Perform a statistical analysis of the dependence of both random variables using **Pearson's** and **Spearman's correlation coefficients**.

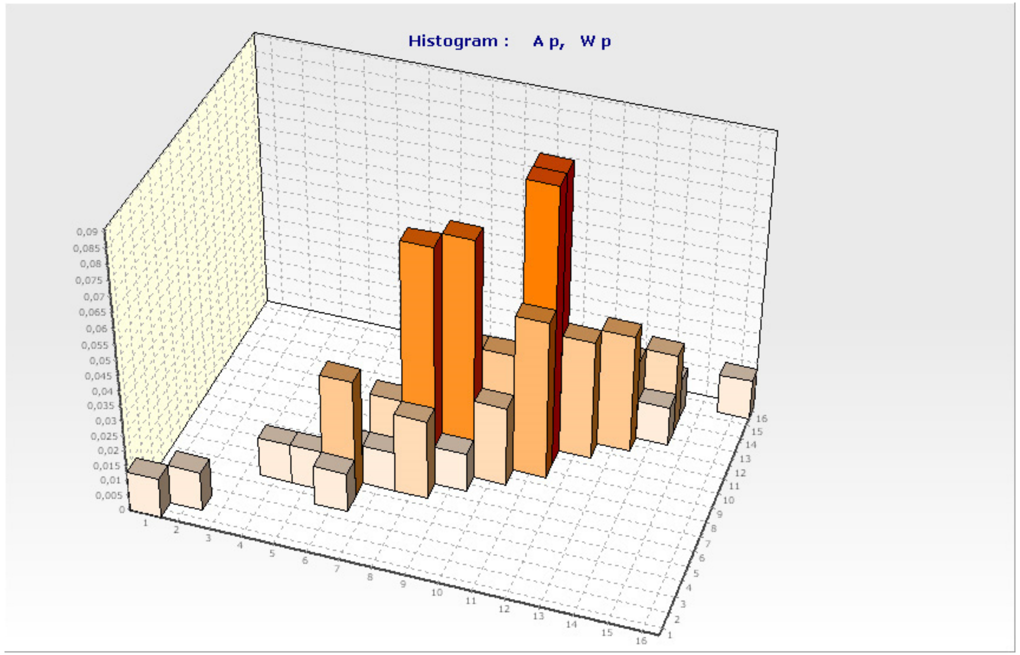
x	y
9	10
7	6
5	1
3	5
1	3



Statistically Dependent Input Variables

Statistically independent random variables are entered into probabilistic calculation using **double** or **triple histograms**.

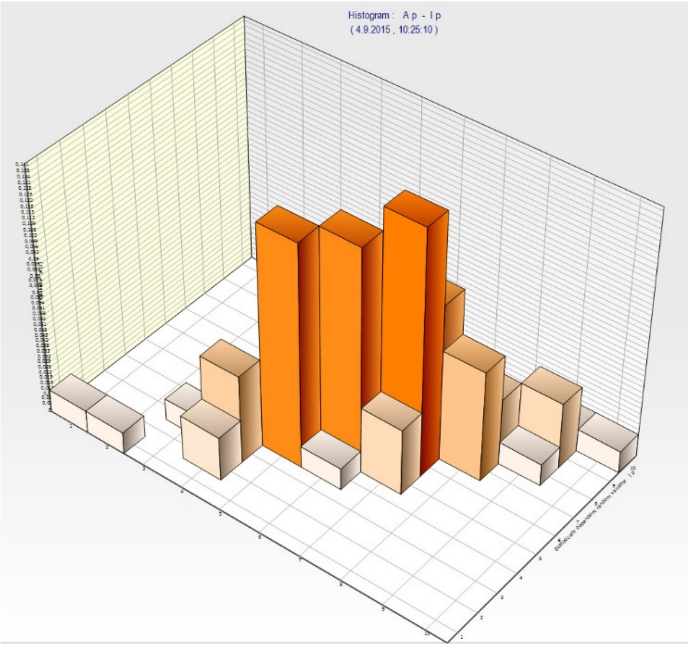




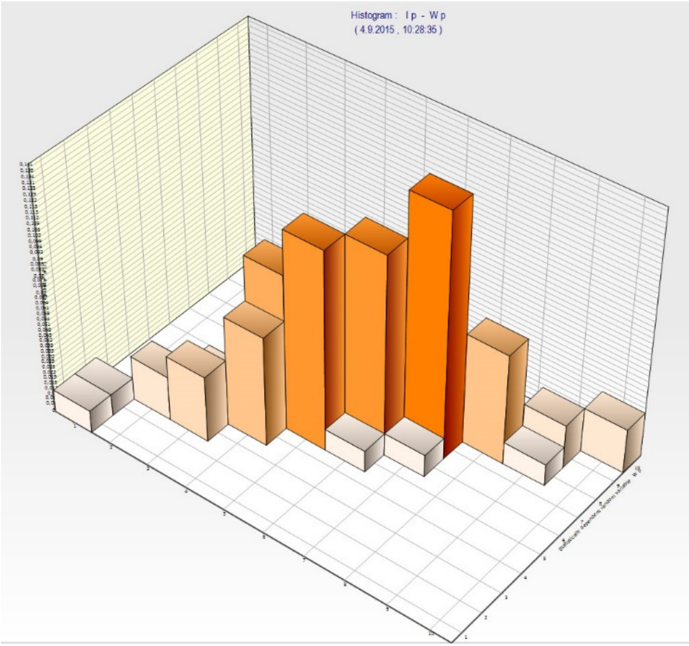
Desktop of HistAn2D: double histogram of statistically independent (left) and dependent (right) random variable

Statistically Dependent Input Variables

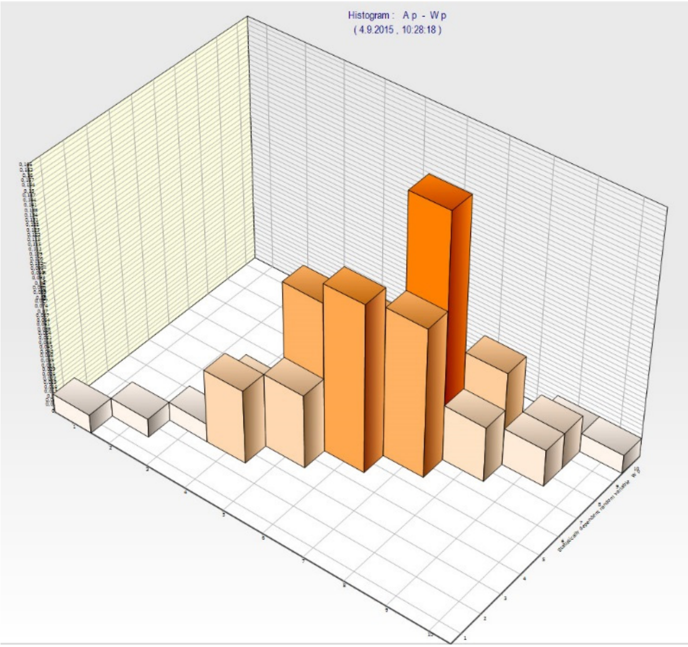
Used double histograms for statistically dependent random **cross-section properties of HE300B** profile.



$A_{var}, I_{y,var}$



$I_{y,var}, W_{y,var}$



$A_{var}, W_{y,var}$

Statistically Dependent Input Variables

Theoretical Background: In each standard histogram A, one axis includes the a_j class which is limited by $a_{\min}$ and $a_{\max}$, while the other axis shows typically the probability, p_{a_j}, of occurrence of that class, a_j.

The sum of probabilities for each class a_j in the histogram is $\sum p_{a_j} = 1$.

In the double histogram of two random variables, Z_1 and Z_2, the quantity z_1 is limited again by $z_{1,\min}$ and $z_{1,\max}$, while z_2 is limited by $z_{2,\min}$ and $z_{2,\max}$.

The values can be divided, using the step $\Delta z_{1,}$ into N_1 intervals for random quantities Z_1, or, using the step Δz_2, into N_2 intervals for the random quantities Z_2. The number of intervals is as follows:

$$N_1 = \frac{z_{1,\max} - z_{1,\min}}{\Delta z_1} \quad \text{and} \quad N_2 = \frac{z_{2,\max} - z_{2,\min}}{\Delta z_2}.$$

Statistically Dependent Input Variables

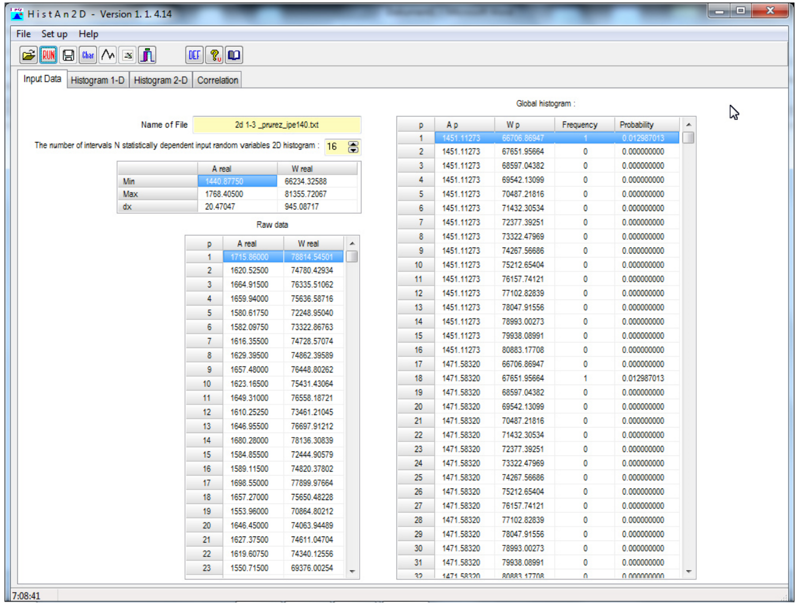
Theoretical Background: If the input variable z_1 is in the j^{th} class of $z_{1,j}$ in theory, z_2 could acquire following values: $z_{2,1}, z_{2,2}, \cdots, z_{2,j}, \cdots, z_{2,N_2}$. This means, it can acquire N_2 values.

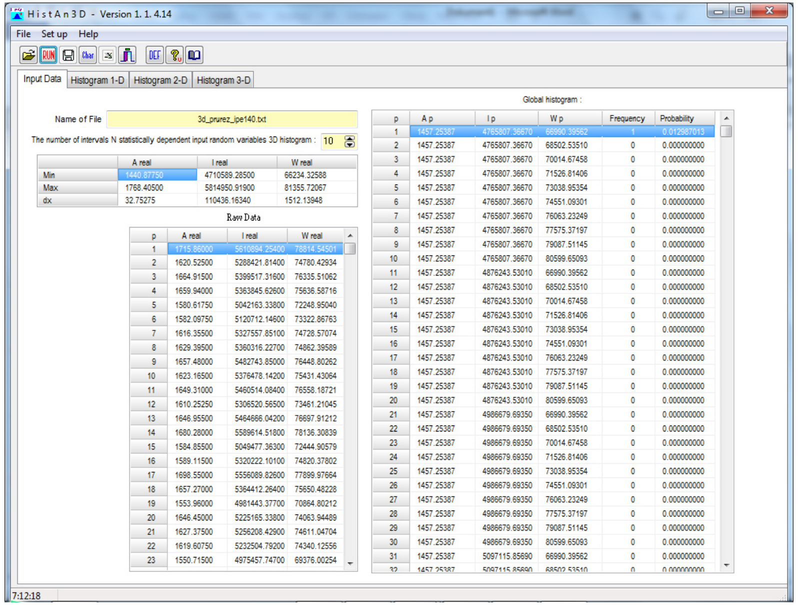
The double histogram of the random quantities z_1 and z_2 can contain $N_1 \cdot N_2$ classes. This means, each class is determined by two values, $z_{1,j}$ and $z_{2,j}$, and by the probability of occurrence of that class, $p_{z_{1,j},z_{2,j}}$. Again: $\sum p_{z_{1,j},z_{2,j}} = 1$.

The number of classes with the non-zero probability can reach the product of $N_1 \cdot N_2$. If the random quantities are dependent, the number of classes in the histogram with the non-zero probability can be considerably lower than the product $N_1 \cdot N_2$.

Software: HistAn2D and HistAn3D

Special software applications **HistAn2D** (left) and **HistAn3D** (right) were developed for creation of the **double** and **triple histograms** which describe the **statistical dependence** between two or three random variables.

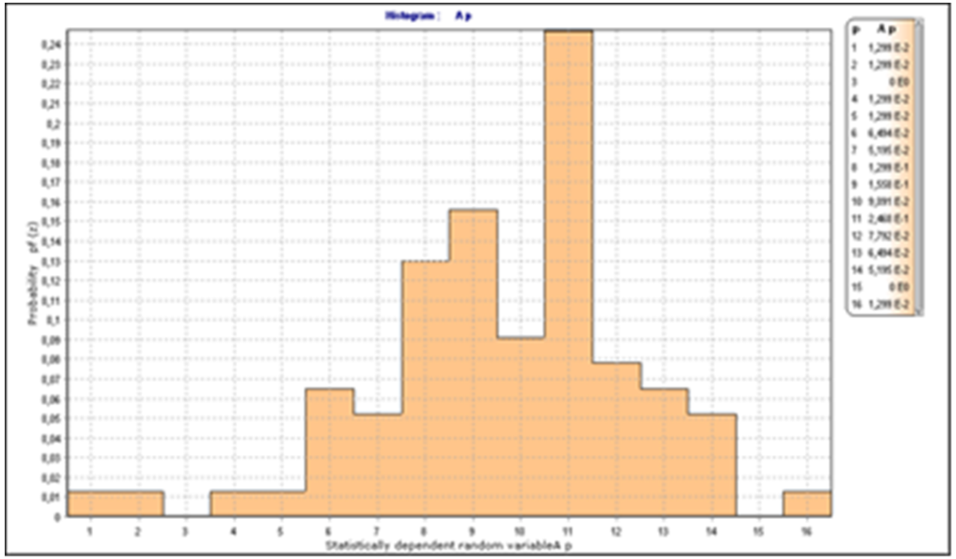


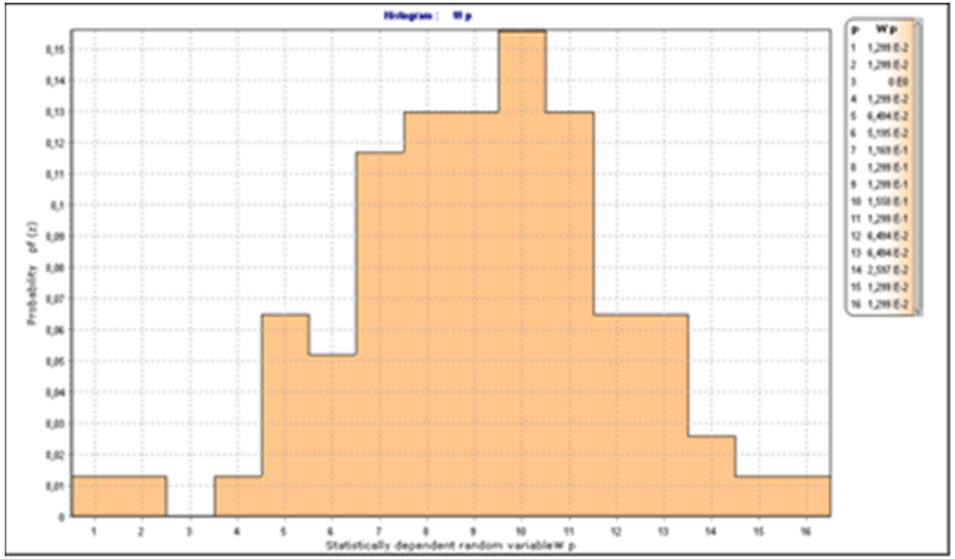


Desktop of **HistAn2D** (left) and **HistAn3D** (right):
raw data of rolled shape IPE 140 cross-section properties under analyses

Software: HistAn2D and HistAn3D

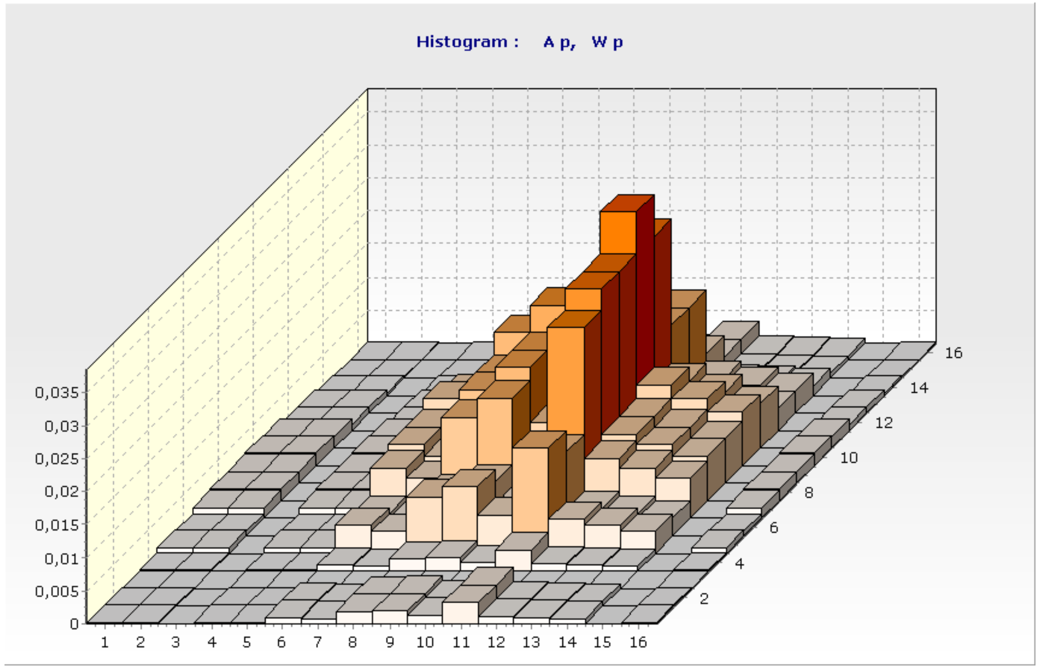
Using the software, it is possible to view for each random variable a simple histogram with non-parametric (empirical) distribution of probability as well as a multidimensional histogram which describes the statistical dependence between the quantities.



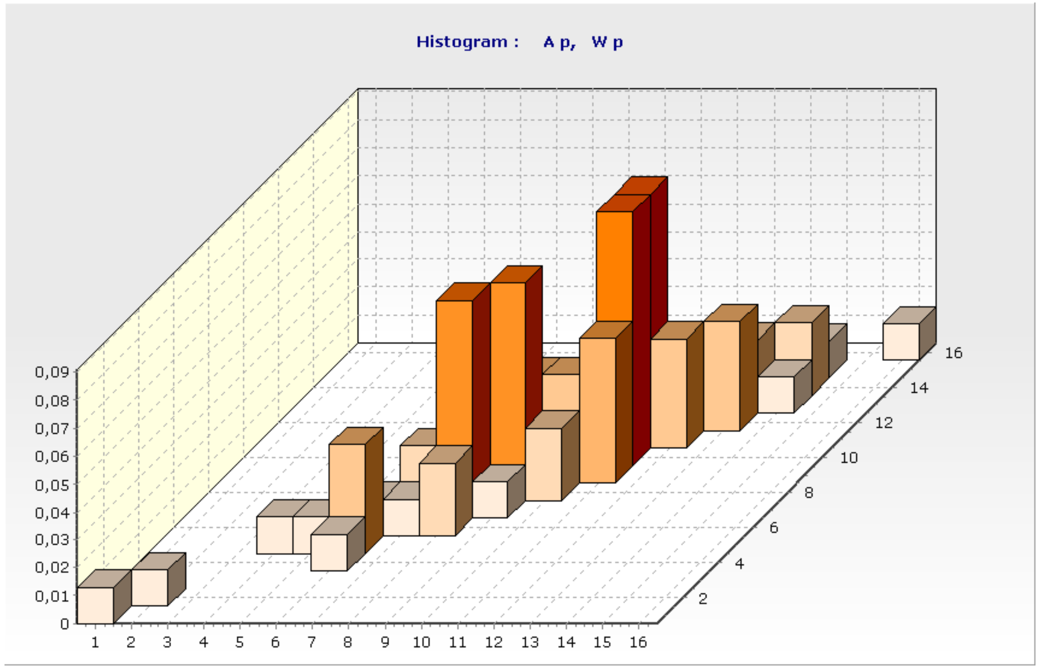


Histograms with **non-parametric** (empirical) **distribution of probability**:
Histogram of the IPE140 cross-section area A (left) and cross-section modulus W_y (right)

Software: HistAn2D and HistAn3D



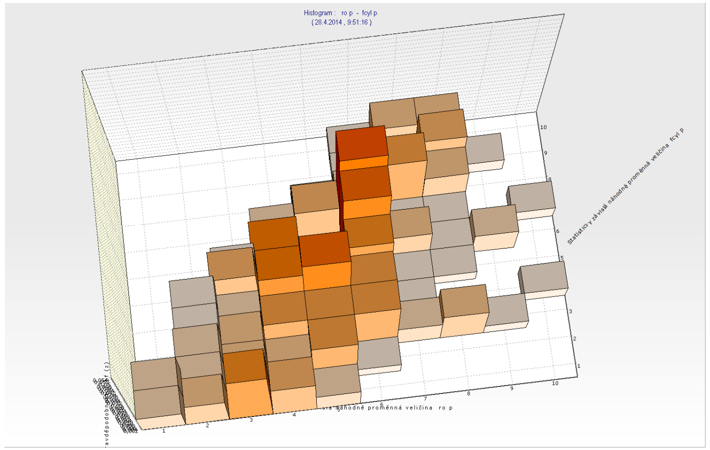
Double histogram for two **statistically independent** random quantities – cross-section area A and cross-section modulus W_y



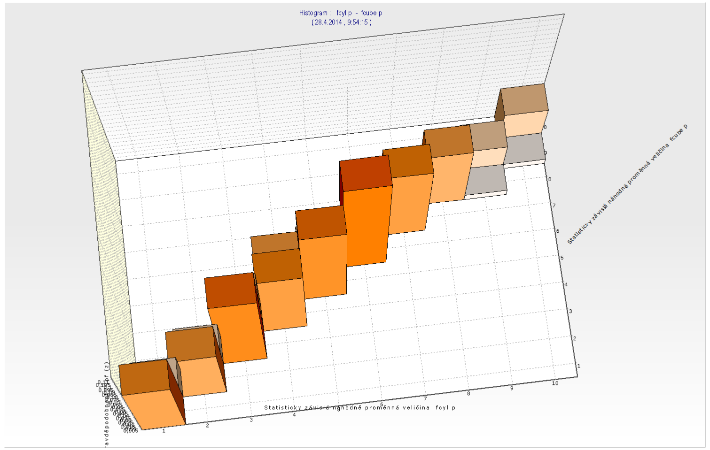
Double histogram for two **statistically dependent** random quantities – cross-section area A and cross-section modulus W_y

Software: HistAn2D and HistAn3D

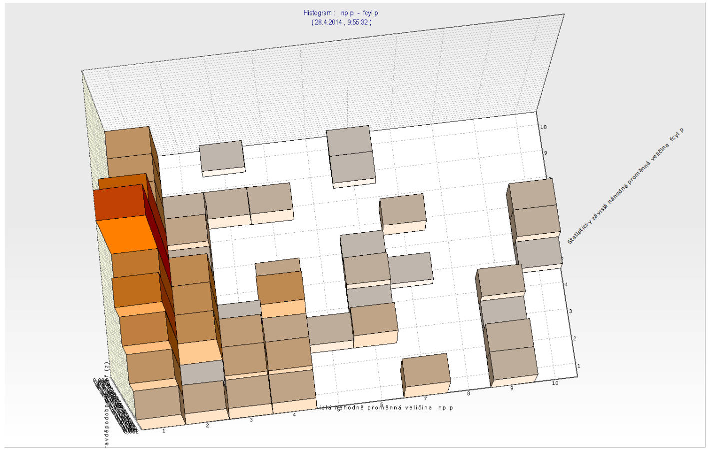
Statistically independent random variables are entered into probabilistic calculation using ProbCalc software



bulk density vs. compressive strength
the correlation 60.8% to 62.2%



cube vs. cylinder compressive strength
the correlation 99.8% to 100.0%

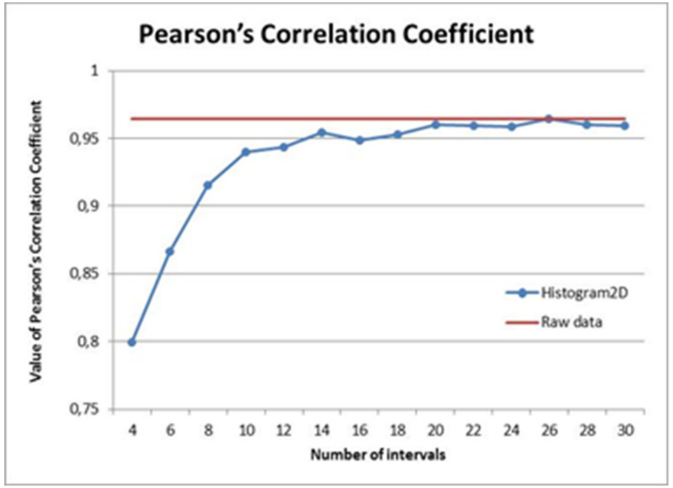


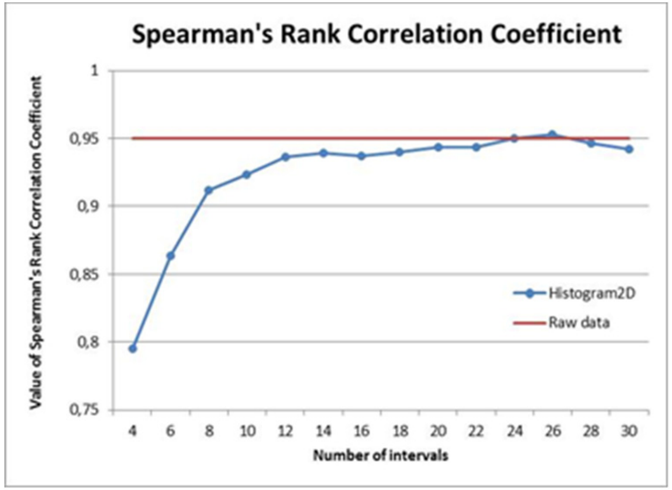
compressive strength of concrete vs. floor in the building
the correlation -21.1% to -25.8%

Software: HistAn2D and HistAn3D

Correlation coefficients of a **double histogram of the statistically dependent quantities** with different numbers of intervals (Pearson’s correlation coefficient for raw data is 0.9645; Spearman correlation coefficient for raw data is 0.9499)

Number of intervals in a double histogram	Pearson’s correlation coefficient	Spearman's rank correlation coefficient	Number of intervals in a double histogram	Pearson’s correlation coefficient	Spearman's rank correlation coefficient
$4^2 = 16$	0.79985097	0.79507798	$18^2 = 324$	0.95267109	0.94023800
$6^2 = 36$	0.86661900	0.86360377	$20^2 = 400$	0.96046634	0.94378886
$8^2 = 64$	0.91530000	0.91194405	$22^2 = 484$	0.95940904	0.94355084
$10^2 = 100$	0.93984931	0.92352904	$24^2 = 576$	0.95903334	0.94989866
$12^2 = 144$	0.94381175	0.93613068	$26^2 = 676$	0.96464064	0.95260826
$14^2 = 196$	0.95443331	0.93939308	$28^2 = 784$	0.96017017	0.94660574
$16^2 = 256$	0.94876401	0.93694950	$30^2 = 900$	0.95938019	0.94245225



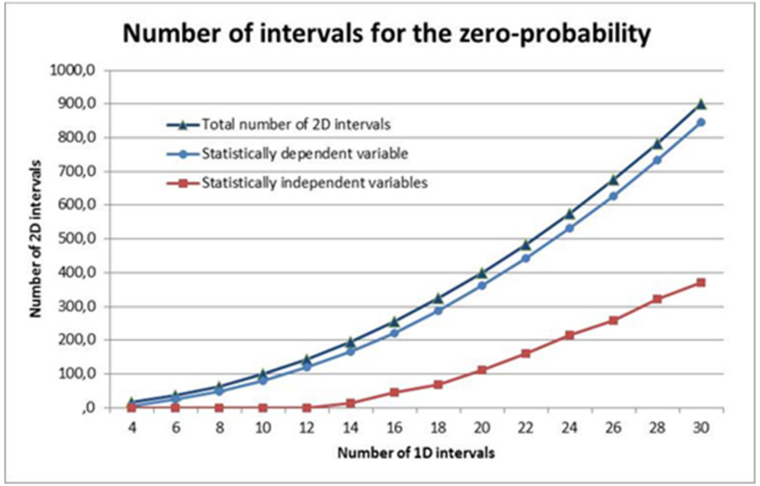


Pearson’s correlation coefficient (up) and **Spearman's rank correlation coefficient** (bottom) of a double histogram vs. number of intervals

Software: HistAn2D and HistAn3D

Number of intervals in a double histogram	Number of zero-probability intervals		Number of intervals in a double histogram	Number of zero-probability intervals	
	Statistically dependent quantities	Statistically independent quantities		Statistically dependent quantities	Statistically independent quantities
$4^2 = 16$	6	0	$18^2 = 324$	288	69
$6^2 = 36$	24	0	$20^2 = 400$	361	112
$8^2 = 64$	48	0	$22^2 = 484$	443	160
$10^2 = 100$	80	0	$24^2 = 576$	531	216
$12^2 = 144$	119	0	$26^2 = 676$	627	258
$14^2 = 196$	166	14	$28^2 = 784$	735	322
$16^2 = 256$	222	46	$30^2 = 900$	847	372

The **number of classes** for double histograms **with zero probability** vs. the number of intervals chosen during creation of the histograms from the primary data



Number of intervals for the zero-probability in double histogram

Software: HistAn2D and HistAn3D

Numerical correlation index – can characterize the dependence between random variables not only for the linear relationship between two variables, but also for nonlinear dependence, or even for more than two random variables:

$$I_k = \frac{T_M - T_C}{T_M}$$

where T_M is the number of all classes in double or triple histogram (for optimal number of intevals and raw data), T_c is the number of non-zero probability classes in double or triple histogram.

For **statistically dependent variables**:

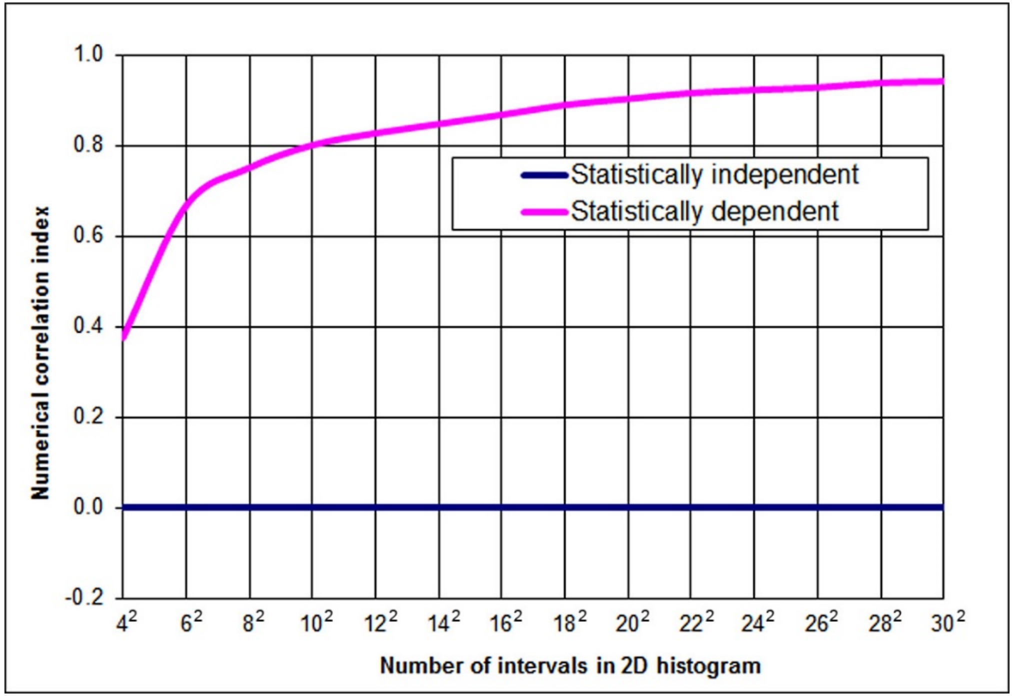
Correction for insufficient number of data:

where $n_1, n_2, n_3, \cdots, n_t$ are the numbers of intervals in histograms, $p_1, p_2, p_3, \cdots, p_t$ are the numbers of intervals without raw data.

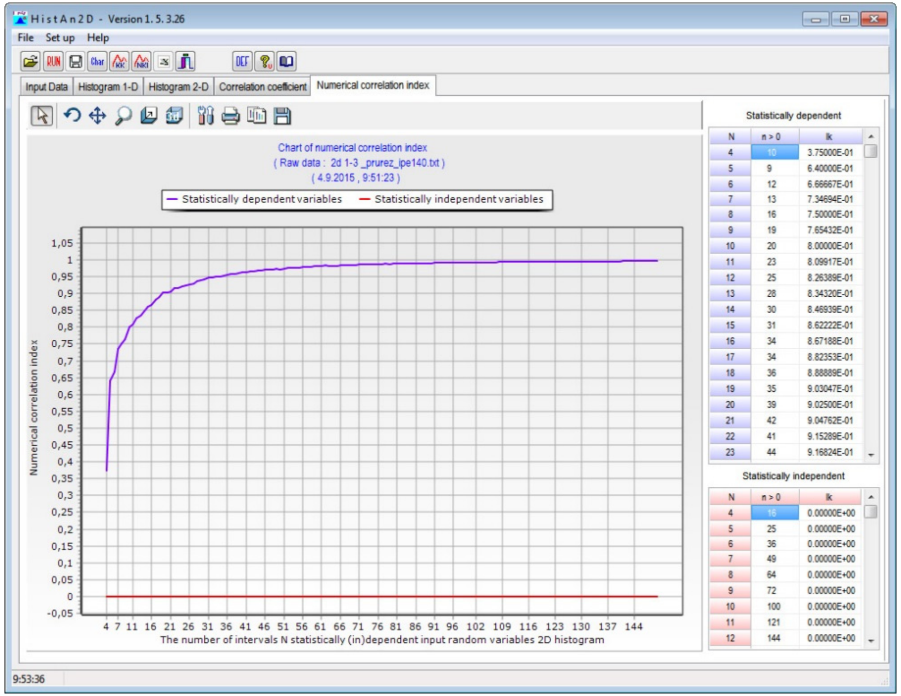
2 dependent variables: $T_M = (n_1 - p_1) \cdot (n_2 - p_2)$

t dependent variables: $T_M = (n_1 - p_1) \cdot (n_2 - p_2) \cdot (n_3 - p_3) \cdot \cdots \cdot (n_t - p_t)$

Software: HistAn2D and HistAn3D

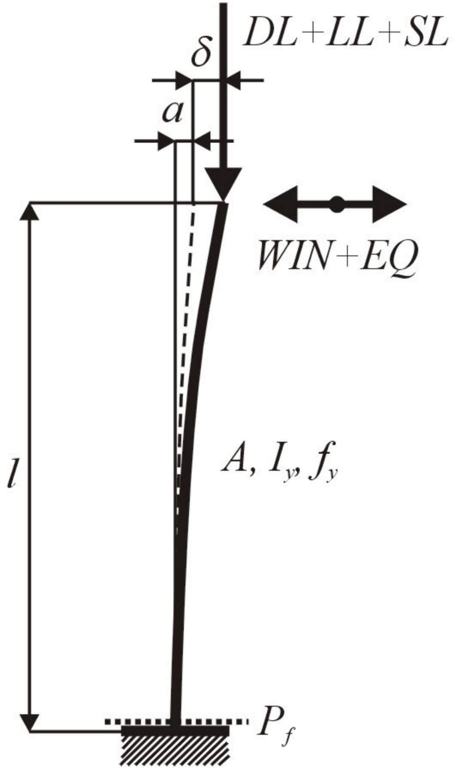


The **numerical correlation index** for two random variables - cross-sectional area A and cross-section modulus W_y



The calculation of **numerical correlation index** in HistAn2D software for variable number of intervals in double histogram

Example 2, Reliability Assessment



Scheme of the structure under assessment

Reliability assessment of the column

l … 6 m
profile HEB 300, **steel S235**, E … $2.1 \cdot 10^{11}$ Pa
imperfections: a … ± 30 mm

Load	Type	Extremal value [kN]
D	Dead	350
L	Long Lasting	75
S	Short Lasting	75
W	Wind	40
EQ	Earthquake	$\frac{1}{20} \cdot (D + L + S) = \frac{500}{20} = 25$

Example 2, Reliability Assessment

Ultimate limit state

$$RF = R - E$$

R … structural resistance – yield stress f_y

E … load effect – stress in outer fibres σ

Serviceability limit state

$$RF = \delta_{tol} - |\delta|$$

δ_{tol} … structural resistance – allowed deformation (35 mm)

δ … load effect – maximal horizontal deformations

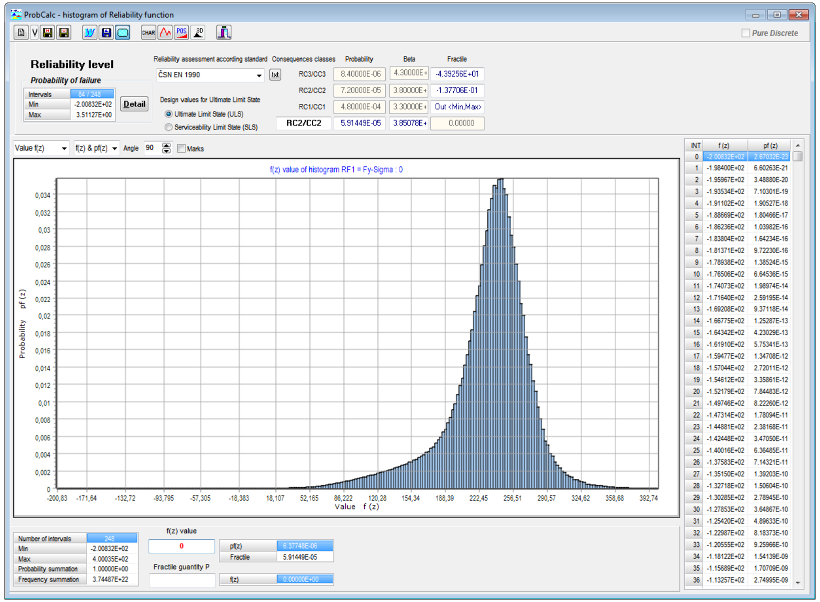
Random input variables:

- 5 load components,
- cross-section variability,
- initial imperfection in column,
- yield stress f_y.

8 random input variables in total

Statistically Dependent Input Variables

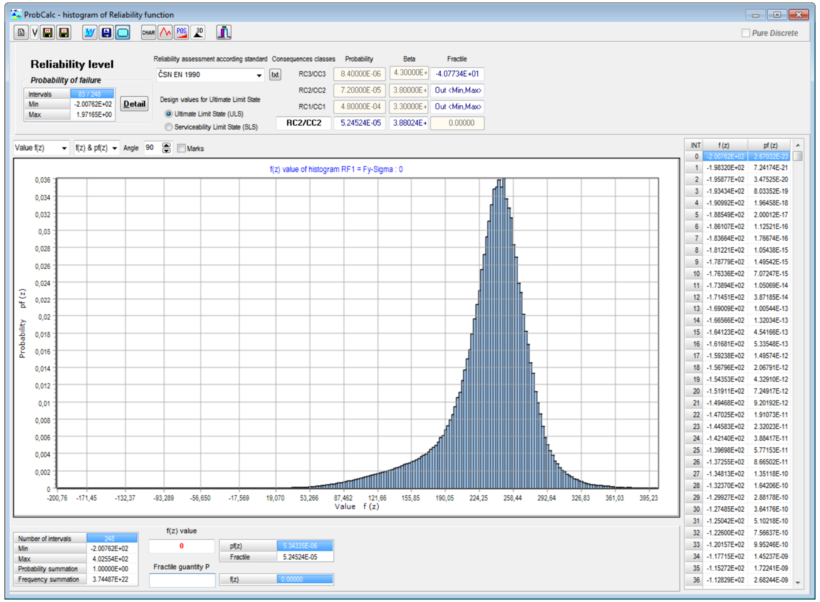
Histograms of reliability function RF, ultimate limit state



Statistically independent cross-section parameters

Failure probability $P_f = 5.133 \cdot 10^{-5}$ (RC2/CC2)

Time of calculation 3:20 min.



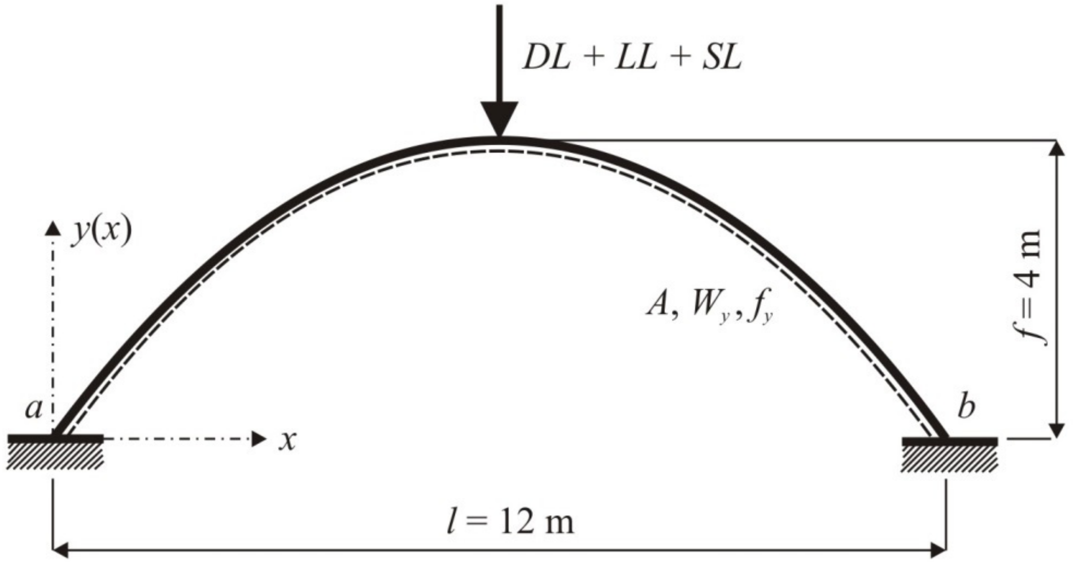
Statistically dependent cross-section parameters

Failure probability $P_f = 5.247 \cdot 10^{-5}$ (RC2/CC2)

Time of calculation 9 sec.

Example 3

Static scheme of the elemental structure of a **parabolic arch** fixed in both ends and loaded with combination of three single loads



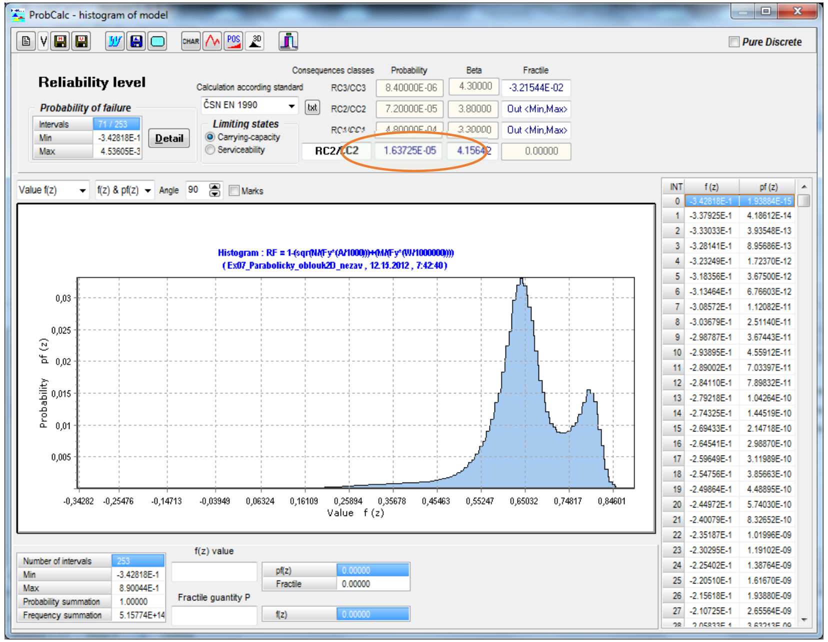
The reliability assessment has been made using the interaction formula:

$$\left(\frac{N_{Ed}}{N_{Rd}}\right)^2 + \frac{M_{Ed}}{M_{Rd}} \leq 1$$

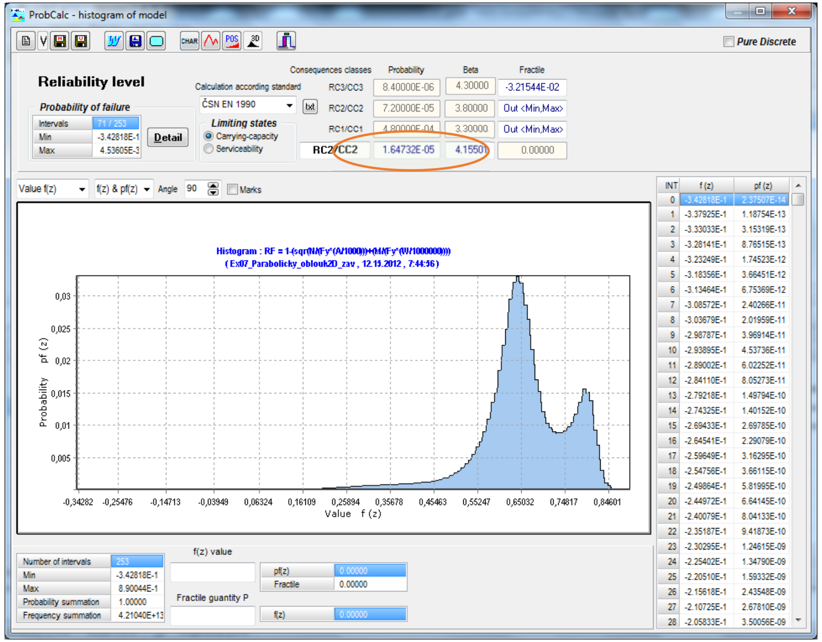
The failure probability P_f was determined using the reliability function RF:

$$P_f = P(RF < 0) = P\left(1 - \left[\left(\frac{N_{Ed}}{N_{Rd}}\right)^2 + \frac{M_{Ed}}{M_{Rd}}\right] < 0\right)$$

Example 3



Histogram of reliability function RF, for the probabilistic calculation with **statistically independent cross-section parameters** of the cross-section area A and cross-section modulus W_y, failure probability $P_f = 1.637 \cdot 10^{-5}$.

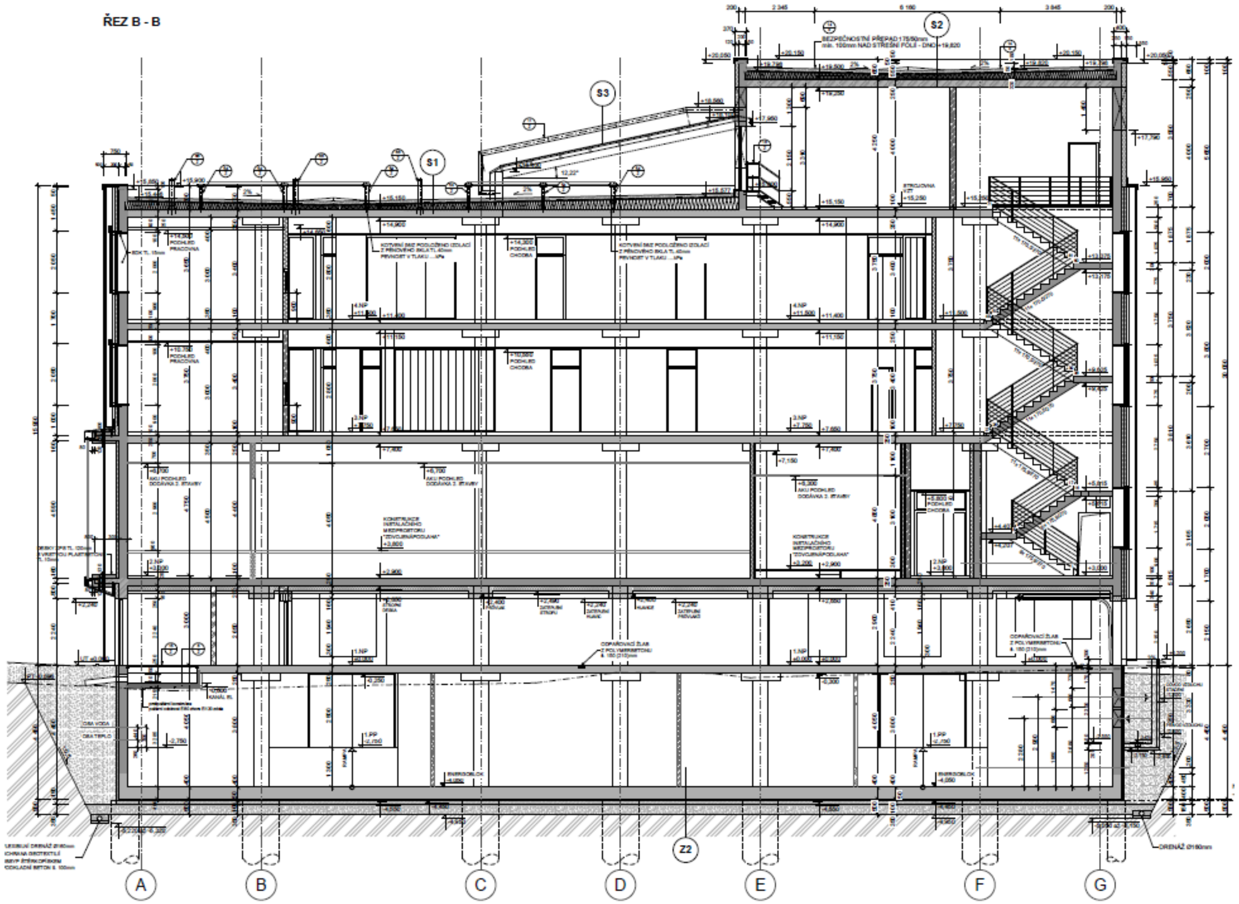


Histogram of reliability function RF, for the probabilistic calculation with statistically dependent cross-section parameters of the cross-section area A and cross-section modulus W_y, failure probability $1.647 \cdot 10^{-5}$.

Structure of Supercomputer Centre

http://www.it4i.cz/

- 5-storey building,
- foundation slab with ribs,
- reinforced concrete structure with walls and columns.



Structure of Supercomputer Centre

http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

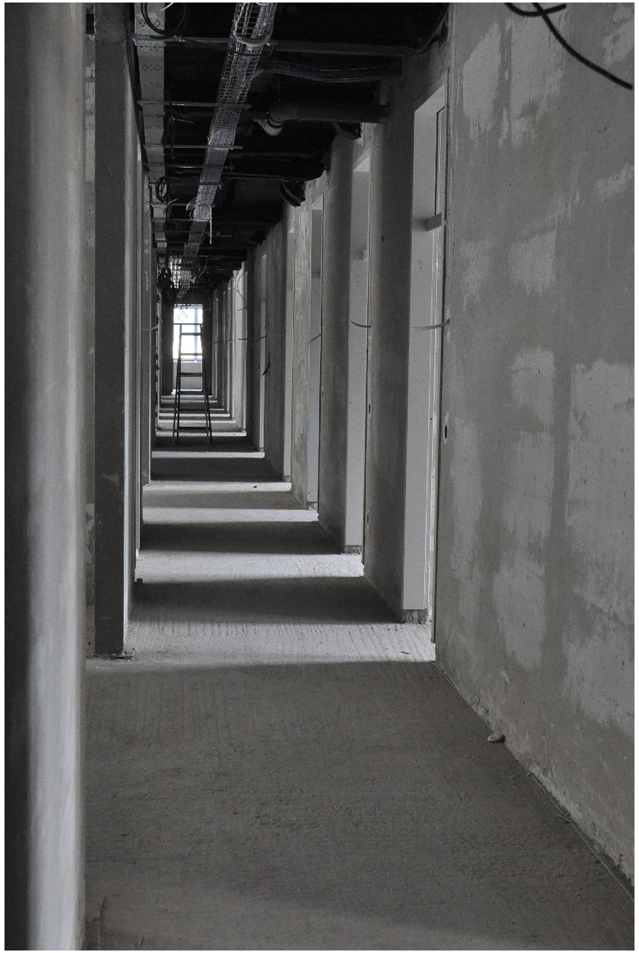
http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

http://www.it4i.cz/

View at construction of the load-carrying structure of the Supercomputing Centre



Structure of Supercomputer Centre

http://www.it4i.cz/

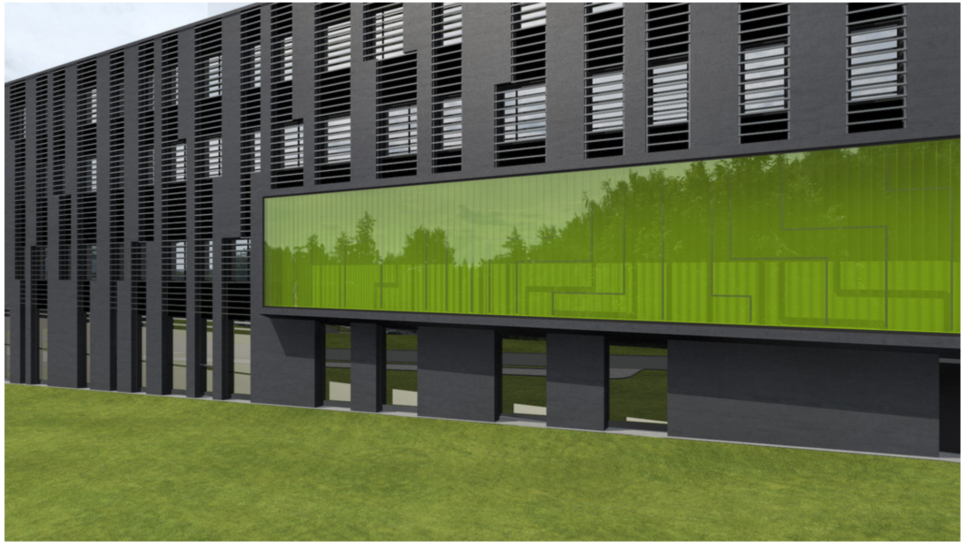
View at construction of the load-carrying structure of the Supercomputing Centre

Structure of Supercomputer Centre

http://www.it4i.cz/

The National Supercomputing Center

IT4 Innovations, March 2014



Foundation Structure of Supercomputer Centre

- Piles foundation under columns,
- Foundation slab with ribs upwards,
- Sliding joint at the bottom surface for volume change,
- Shrink bands elimination.

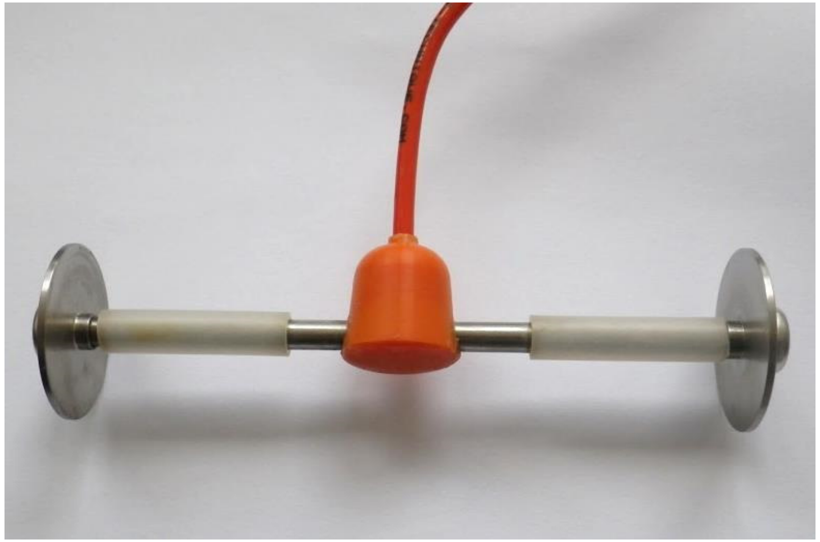
Foundation Structure of Supercomputer Centre

Goal of the measurements:

- Development of the hydration heat during foundation slab concreting,
- Change of the stresses in the concrete,
- Comparison the tensile stresses with the numerical calculation.

Technology of Measurement

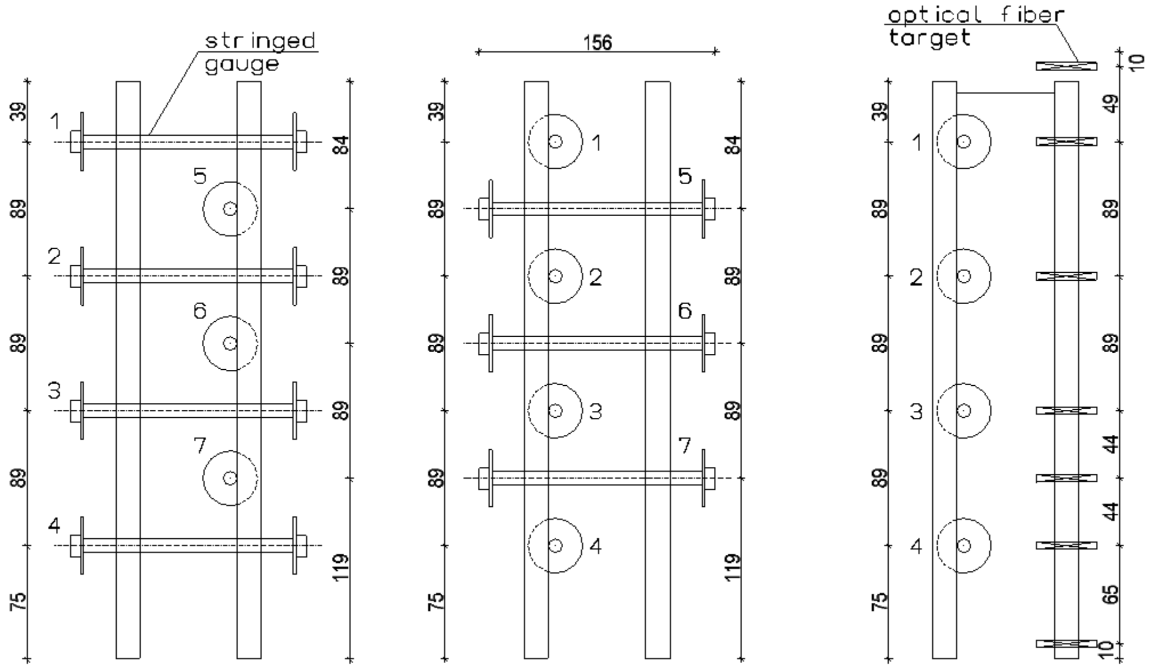
- optical fiber - temperature
- string gauges - temperature/stress
- foil strain gauges - strain (stress)
- digital thermometer - temperature



Setting of Measuring Column

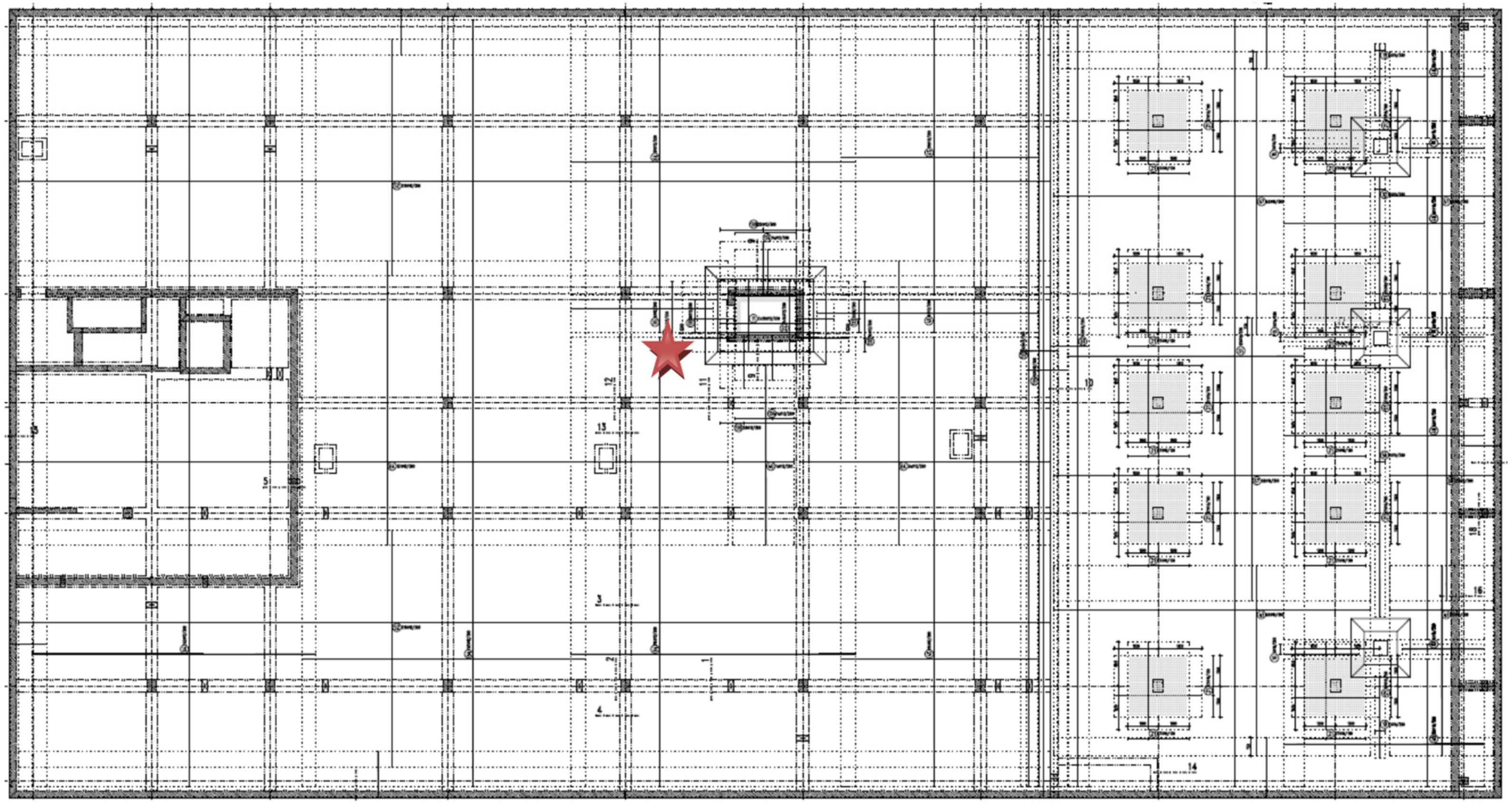
Measuring column mounted to a reinforced concrete slab

- 7x string gauges (3 in transverse and 4 in longitudinal direction)
- 6x bundles of optical fibers



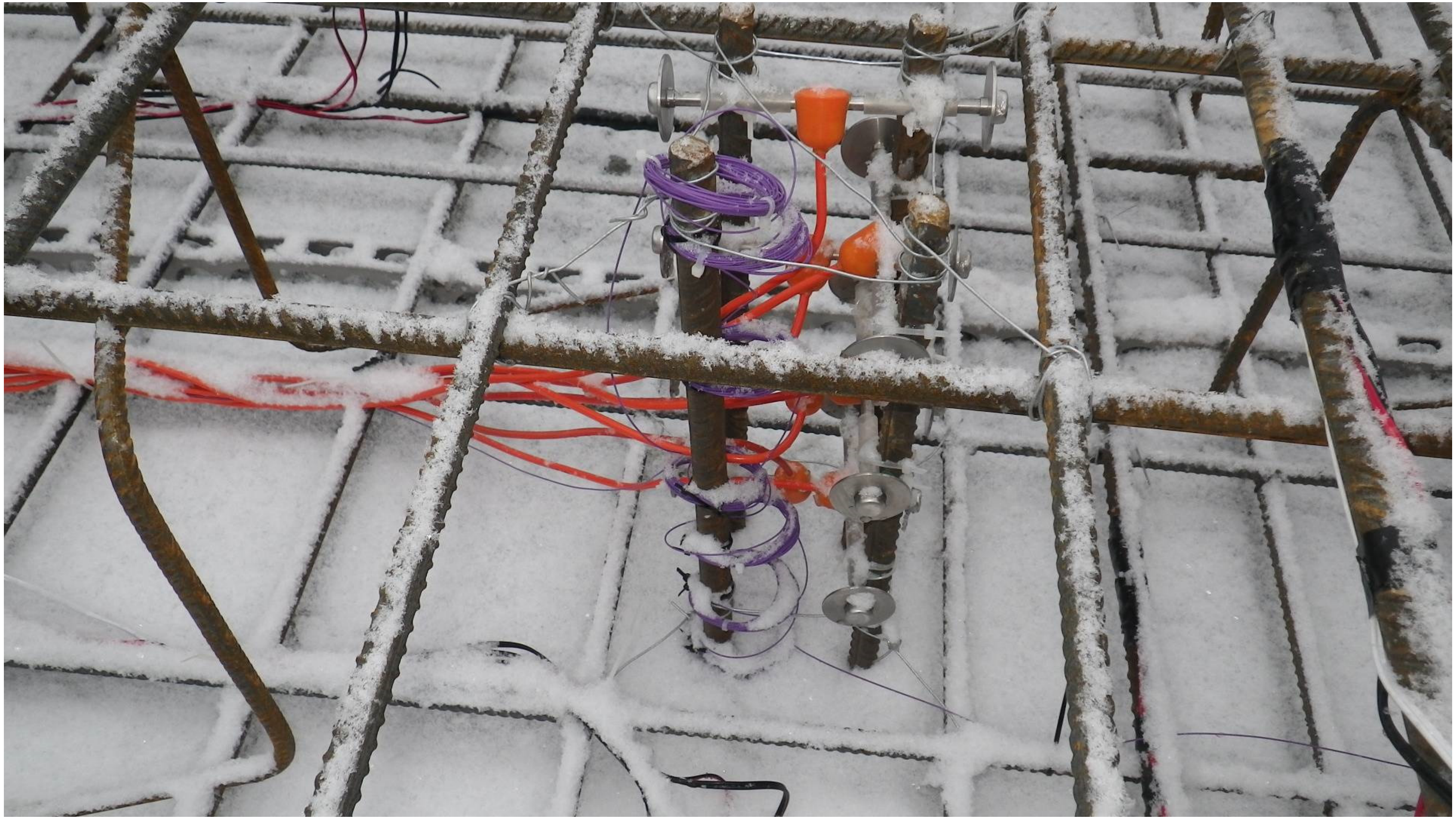
Ground Plan of Foundation Slab

Schematic chart of the foundation slab and location of the sensors



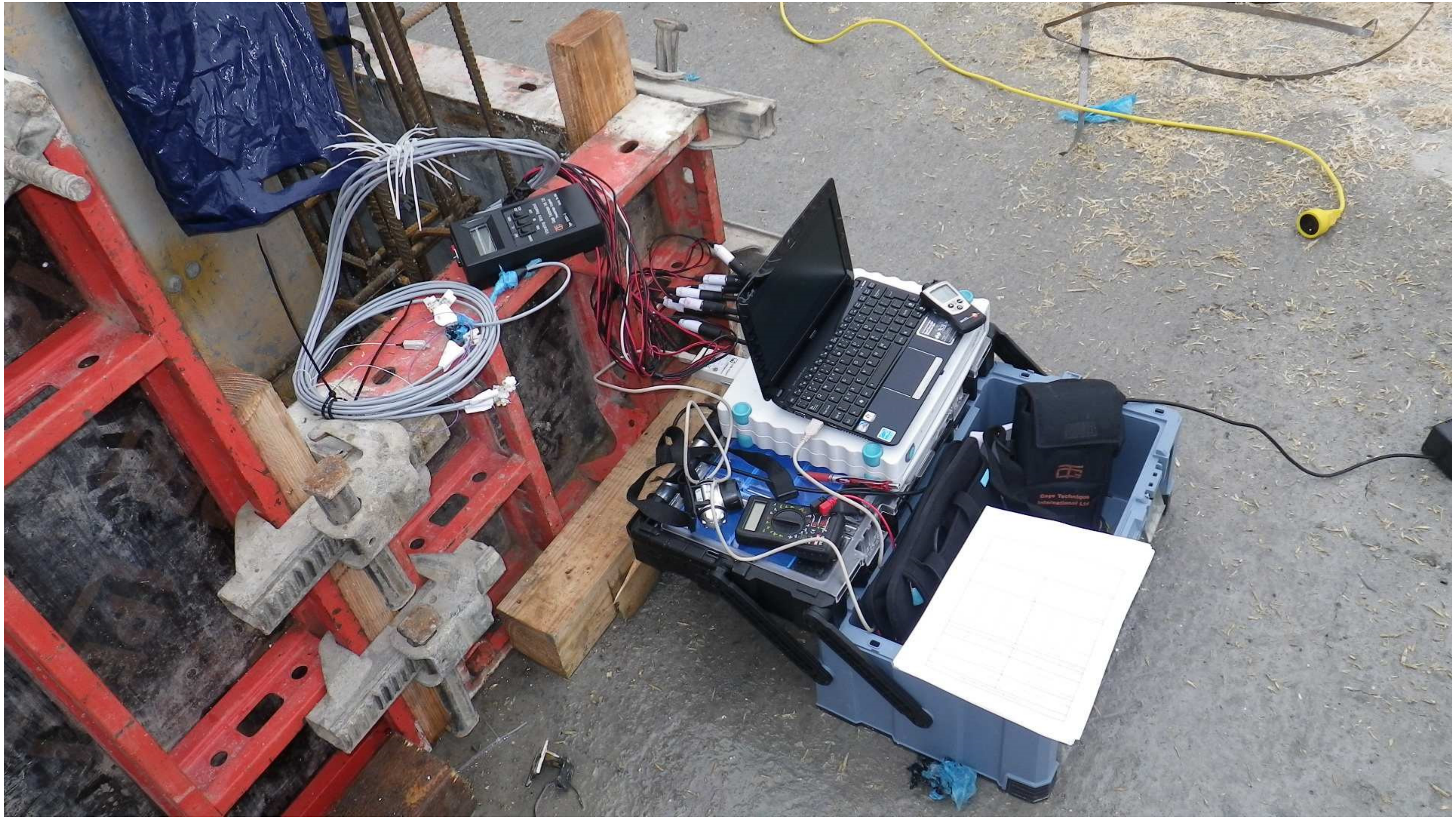
Installation of a Measuring Pillar in the Structure

Installation of a Measuring Pillar in the Structure



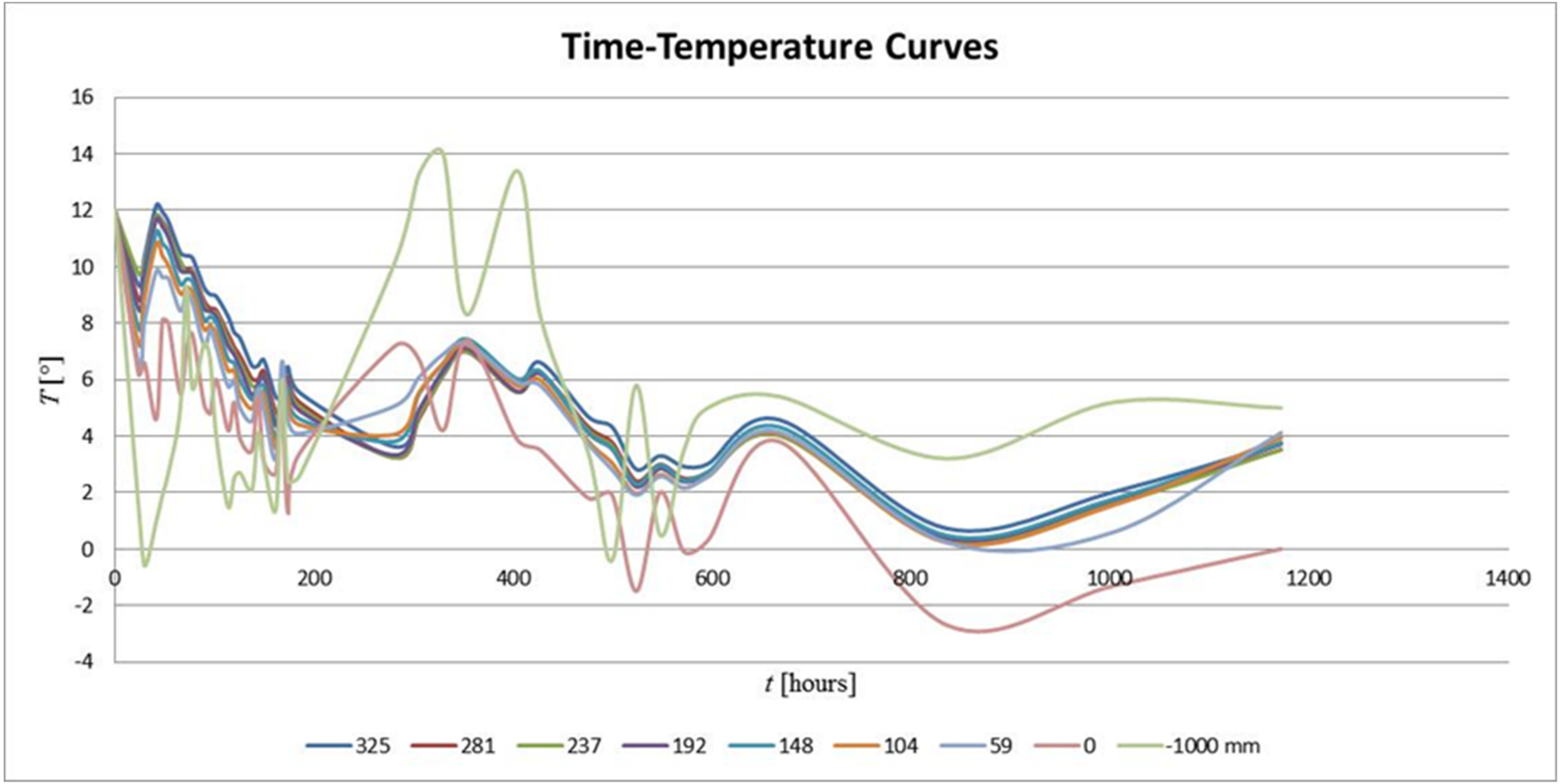
Installation of a Measuring Pillar in the Structure

Data Transfer on Construction Site



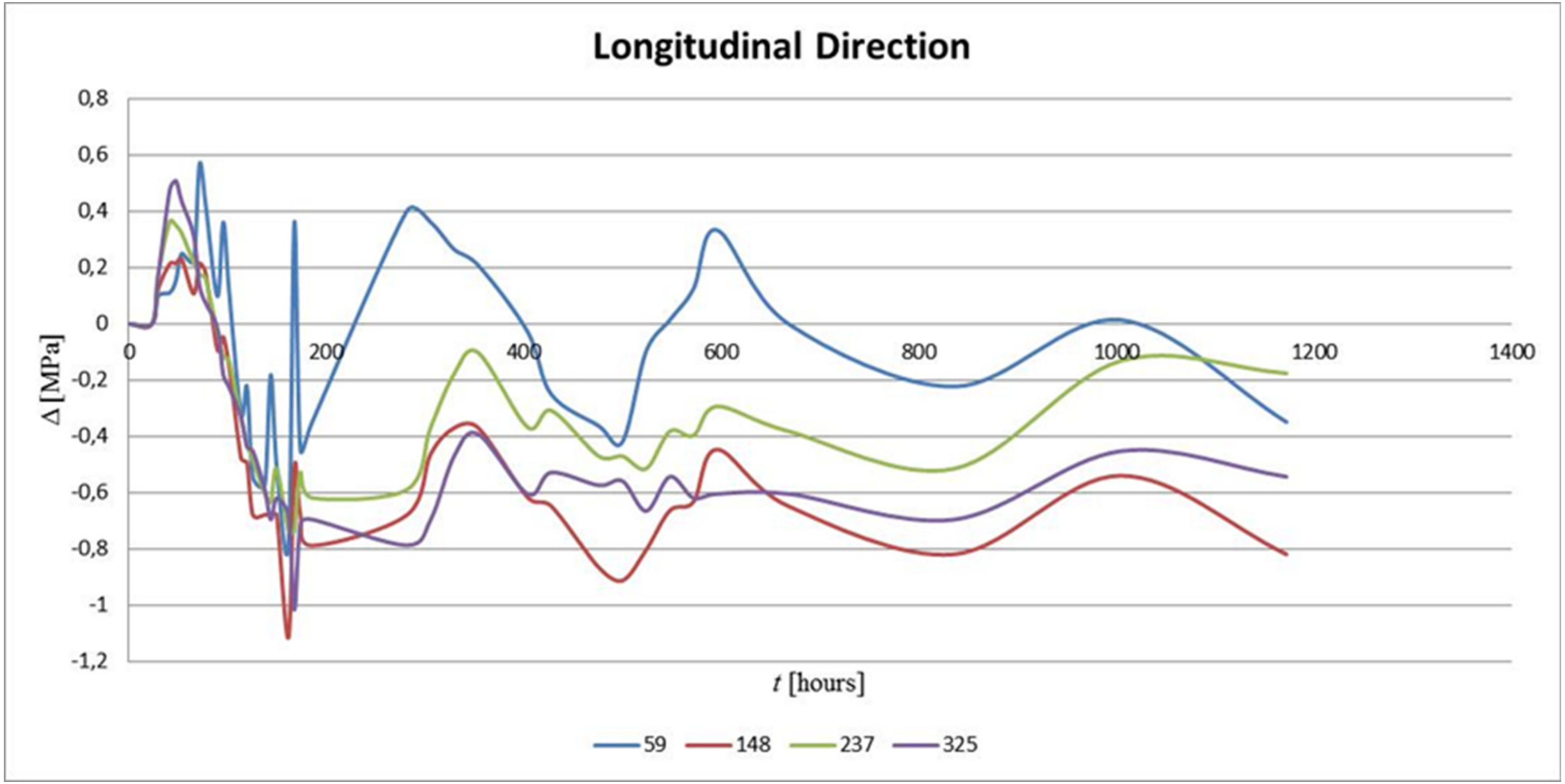
Time Records of the Measured Data

Time records of the temperature in the concrete foundation slab made for 9 different heights



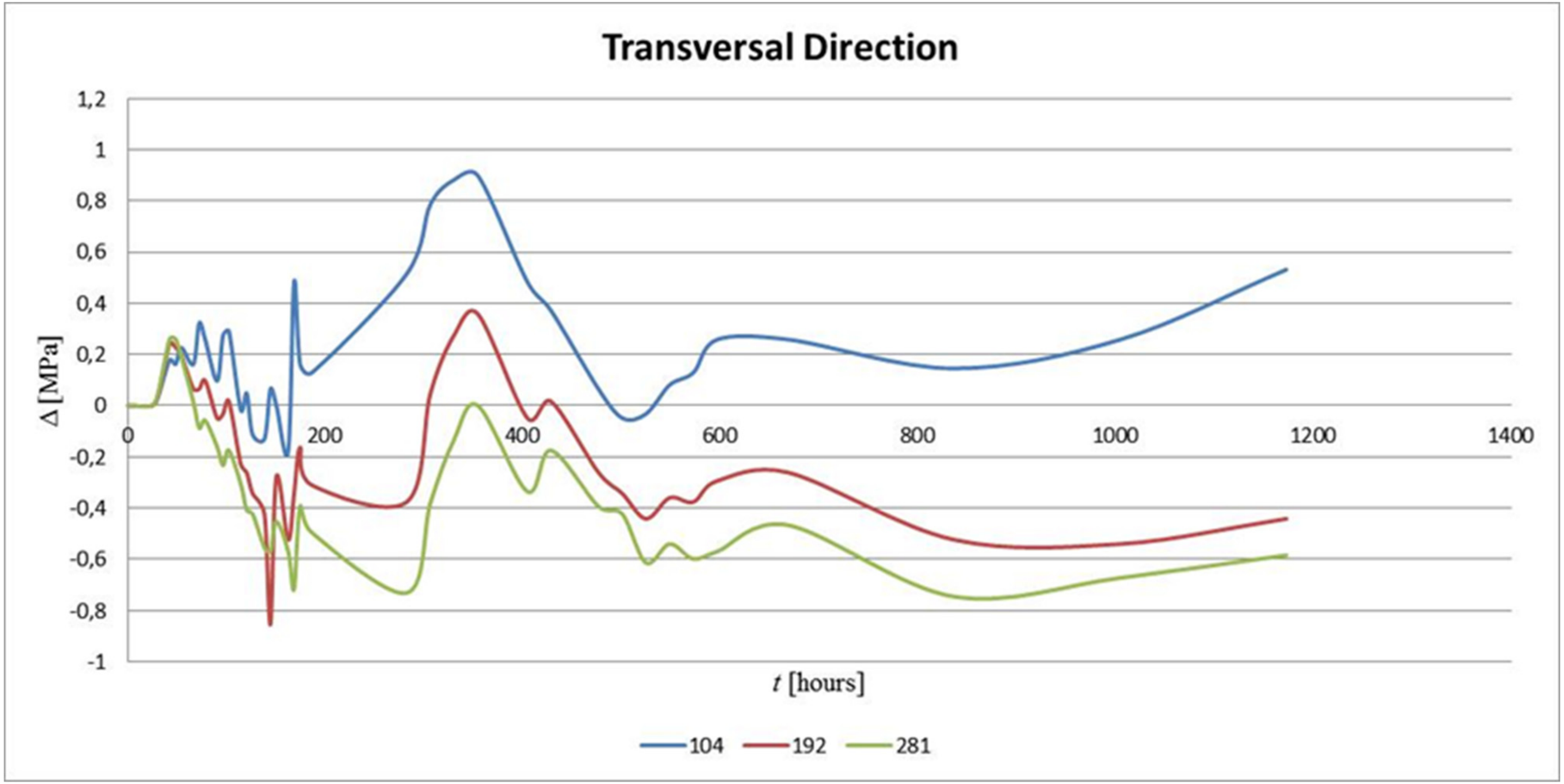
Time Records of the Measured Data

Changes in time of the normal stress in the concrete slab in a longitudinal direction for 4 different heights



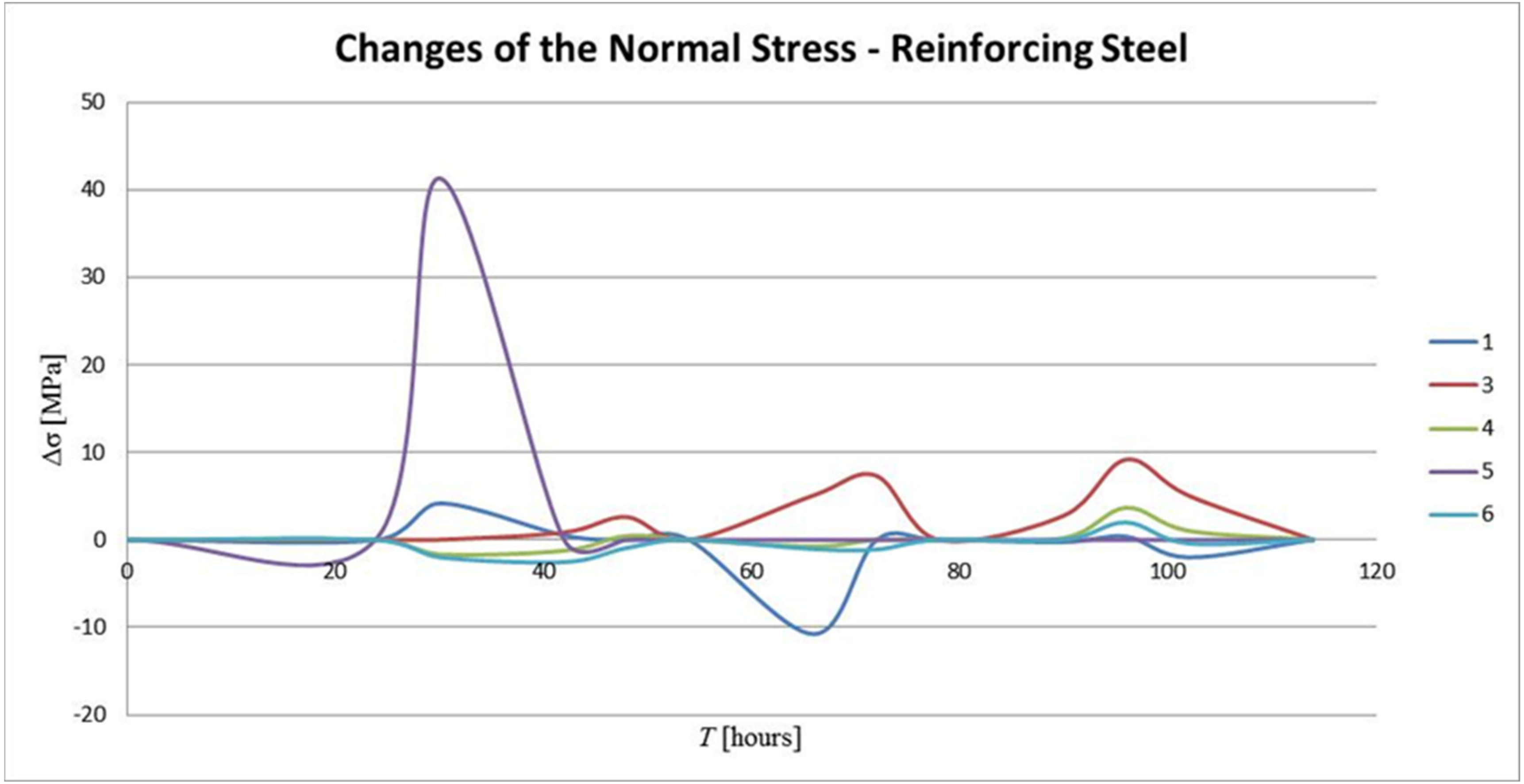
Time Records of the Measured Data

Changes in time of the normal stress in the concrete slab in a transversal direction for 3 different heights



Time Records of the Measured Data

Changes in time of the normal stress in the reinforcing steel in the foundation slab



Analyzing the Temperature in the Foundation Slab

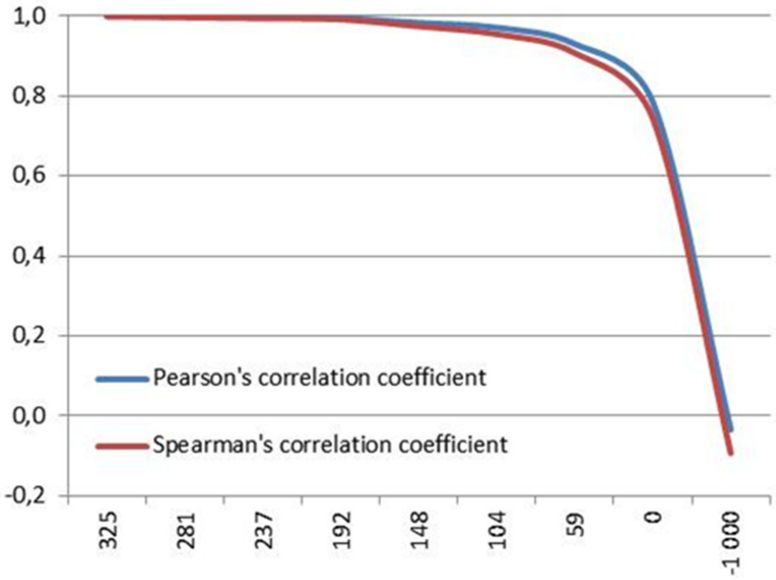
[mm]	325	281	237	192	148	104	59	0	-1000
325	1	0.999179	0.996923	0.995500	0.983759	0.971624	0.930670	0.785600	-0,034170
281		1	0.998172	0.997997	0.988557	0.978113	0.941768	0.793464	-0,129316
237			1	0.997222	0.988365	0.978245	0.942759	0.798301	-0.007715
192				1	0.995390	0.988138	0.957607	0.809621	0.028066
148					1	0.997432	0.978363	0.839512	0.106057
104						1	0.988934	0.859433	0.148671
59							1	0.899881	0.252198
0		sym.						1	0.298942
-1000									1

The correlation matrix which uses the **Pearson's correlation coefficients** to describe the statistics dependence of the temperature at different heights of the foundation slab and ambient temperature

[mm]	325	281	237	192	148	104	59	0	-1000
325	1	0.997100	0.994693	0.992175	0.974777	0.955080	0.907967	0.744175	-0.093153
281		1	0.998468	0.996990	0.981617	0.963234	0.919626	0.749822	-0.725701
237			1	0.998523	0.985556	0.969800	0.928818	0.762627	-0.051773
192				1	0.989439	0.975540	0.938054	0.769048	-0.024796
148					1	0.993216	0.967774	0.800925	0.068958
104						1	0.985939	0.836969	0.115915
59							1	0.881010	0.207761
0		sym.						1	0.244649
-1000									1

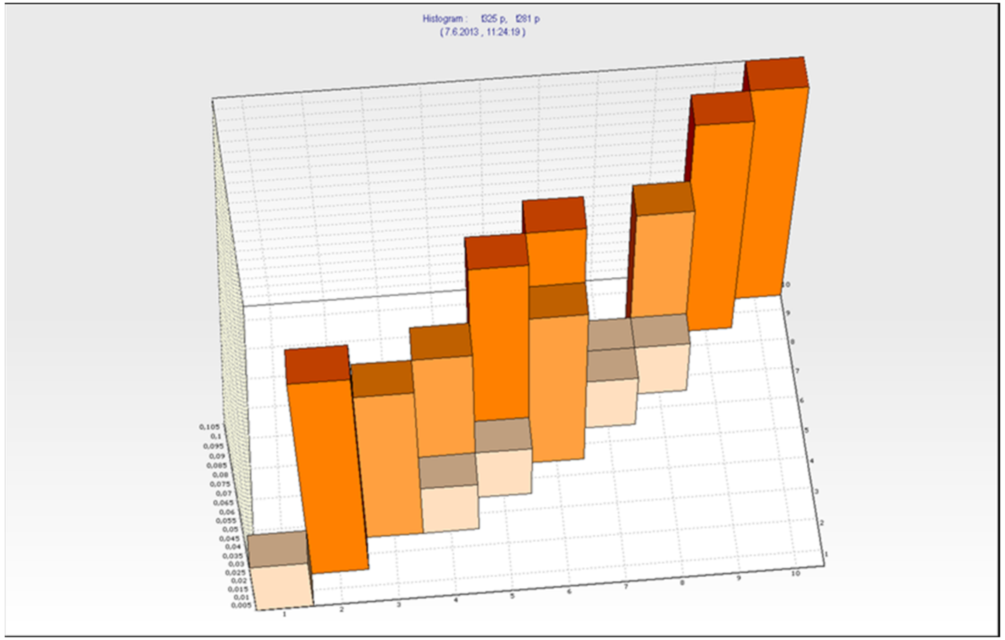
The correlation matrix which uses the **Spearman's correlation coefficients** to describe the statistics dependence of the temperature at different heights of the foundation slab and ambient temperature

Analyzing the Temperature in the Foundation Slab



The **Pearson's** and **Spearman's correlation coefficients** between the temperatures measured at +325 mm and temperatures in other measuring points

Desktop of **HistAn2D**: The double histogram which describes the statistic dependence of the random variable temperature at +325 mm and +281 mm using DOProC method



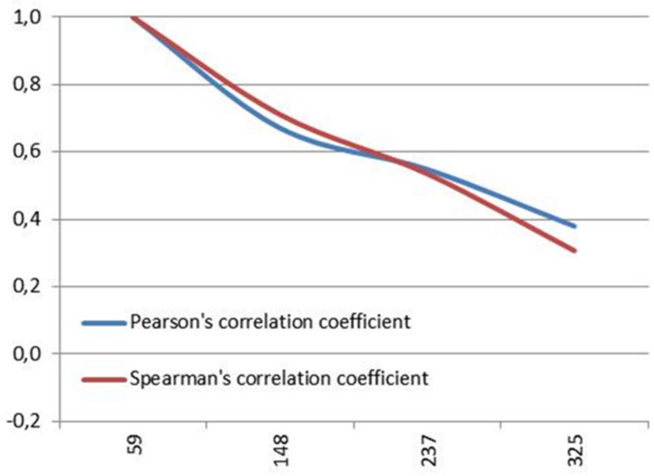
Analyzing Changes in Normal Stress in a Slab

	59	148	237	325
59	1	0.671255	0.548708	0.379817
148		1	0.906740	0.894987
237			1	0.906740
325	sym.			1

	59	148	237	325
59	1	0.711119	0.535419	0.307283
148		1	0.860507	0.784488
237			1	0.860507
325	sym.			1

The correlation matrix which describes the statistic dependence of changes in the **normal stress** in a concrete slab in a **longitudinal direction** in four heights using the **Pearson's** (left) and **Spearman's** (right) **correlation coefficient**

The Pearson's and Spearman's correlation coefficient for the randomly variable change in the **normal stress** of the foundation slab in a **longitudinal direction** measured at +59 mm combined with the values at other heights

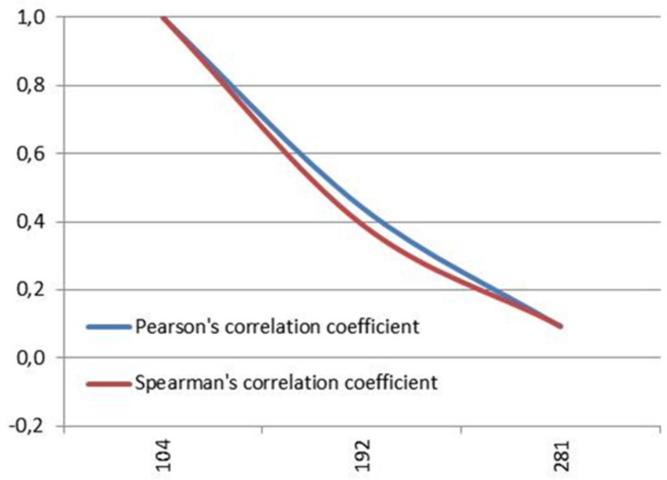


Analyzing Changes in Normal Stress in a Slab

	104	192	281
104	1	0.443257	0.092120
192		1	0.874841
281			1

	104	192	281
104	1	0.393268	0.094875
192		1	0.918306
281			1

The correlation matrix which describes the statistic dependence of changes in the **normal stress** in a concrete slab in a **transversal direction** in three heights using the **Pearson's** (left) and **Spearman's** (right) **correlation coefficient**



The **Pearson's** and **Spearman's correlation coefficient** for the randomly variable change in the **normal stress** of the concrete foundation slab in a **transversal direction** measured at +104 mm combined with the values measured at other heights

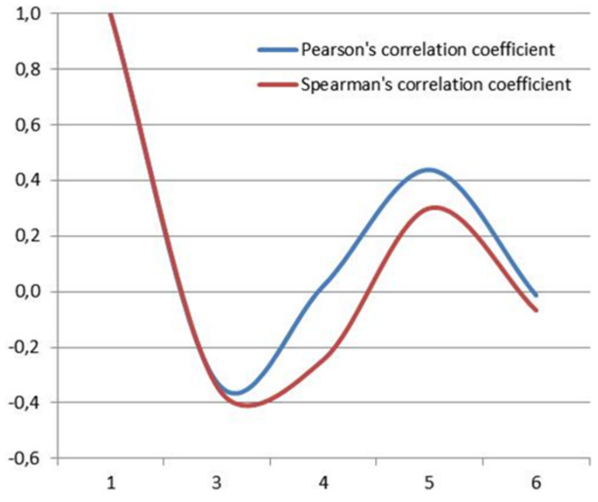
Analyzing Changes in Normal Stress in Reinforcement

sensor	1	3	4	5	6
1	1	-0.326076	0.020921	0.438671	-0.013040
3		1	0.659559	-0.267758	0.357575
4			1	-0.426226	0.861452
5				1	-0.401796
6					1

sensor	1	3	4	5	6
1	1	-0.340521	-0.244565	0.300965	-0.067029
3		1	0.369197	-0.254437	-0.172053
4			1	-0.541736	0.632246
5				1	-0.290020
6					1

The correlation matrix which describes the statistic dependence of changes in the **normal stress** in **reinforcing steel** in the concrete slab using five sensors and the **Pearson's correlation coefficient** (left) and the **Spearman's correlation coefficient** (right)

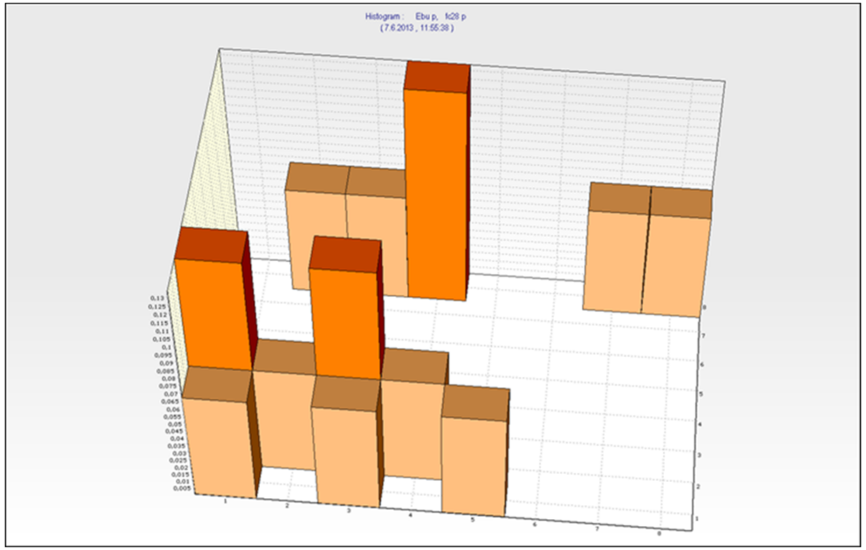
The **Pearson's** and **Spearman's correlation coefficient** for the randomly variable change in the **normal stress** in **reinforcing steel** in the foundation slab measured by the sensor #1 combined with the values measured by the sensors #3 through #6



Statistical Analysis of Dependence

Analyzing the statistic dependence of the **dynamic modulus of elasticity of the concrete** and **compressive cube strength of the concrete**:

- **Dynamic modulus of elasticity** of the concrete and **compressive cube strength** of the concrete - analyze using the **non-destructive tests**.
- The statistic dependence between the two randomly variable quantities can be described again using a pair of correlation coefficients (the **Pearson's** and **Spearman's correlation coefficient** equal to **0.541351** and **0.524191**, respective) or using a double histogram.



Desktop of HistAn2D: A **double histogram** which describes the statistic dependence of the **dynamic modulus of elasticity** of the concrete and **compressive cube strength** of the concrete

Statistical Analysis of Dependence

Statistical analysis of concrete in the frame structure:

- Destructive testing detected physical and mechanical parameters of concrete in supporting structure and the floors during the construction and technical survey of the building of the Faculty of Mechanical Engineering in Brno.
- The aim of this survey was the need to assess the **quality of concrete** in selected parts of the **horizontal and vertical load-bearing structures**.

View on the structure with old cladding during inspection

Statistical Analysis of Dependence

Statistical analysis of concrete in the frame structure

Detail of the concrete frame under cladding before reconstruction

Statistical Analysis of Dependence

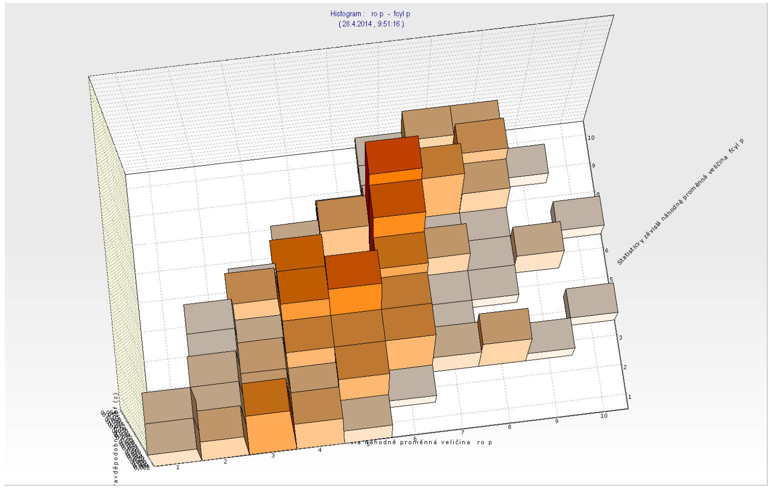
Statistical analysis of concrete in the frame structure:

- 32 cores 45 mm in diameter on the peripheral columns over the entire height of the 17 storied building was done.
- In the internal load-bearing columns between 1st and 10th over ground floors of building were performed 25 cores.
- In the horizontal supporting structures were done 6 cores in total on floors 2, 4 and 6.
- From all of the cores were created 166 test specimens to determine strength parameters of concrete.

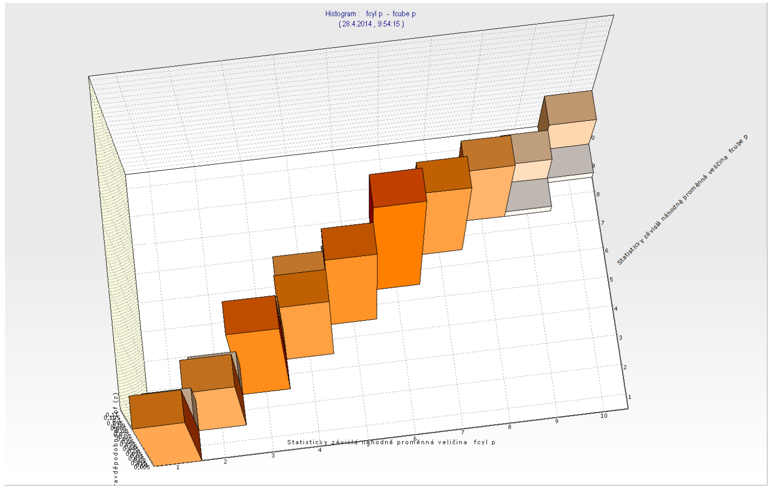
Statistical Analysis of Dependence

Statistical analysis of concrete in the frame structure:

Using the statistical and sensitivity analysis was subsequently found a statistical dependence between the **bulk density of concrete** and **concrete compressive strength** with the correlation in range 60.8% to 62.2% and between the **cube** and **cylinder compressive strength of concrete** with the correlation of 99.8% to 100.0%.



bulk density vs. compressive strength
the correlation 60.8% to 62.2%

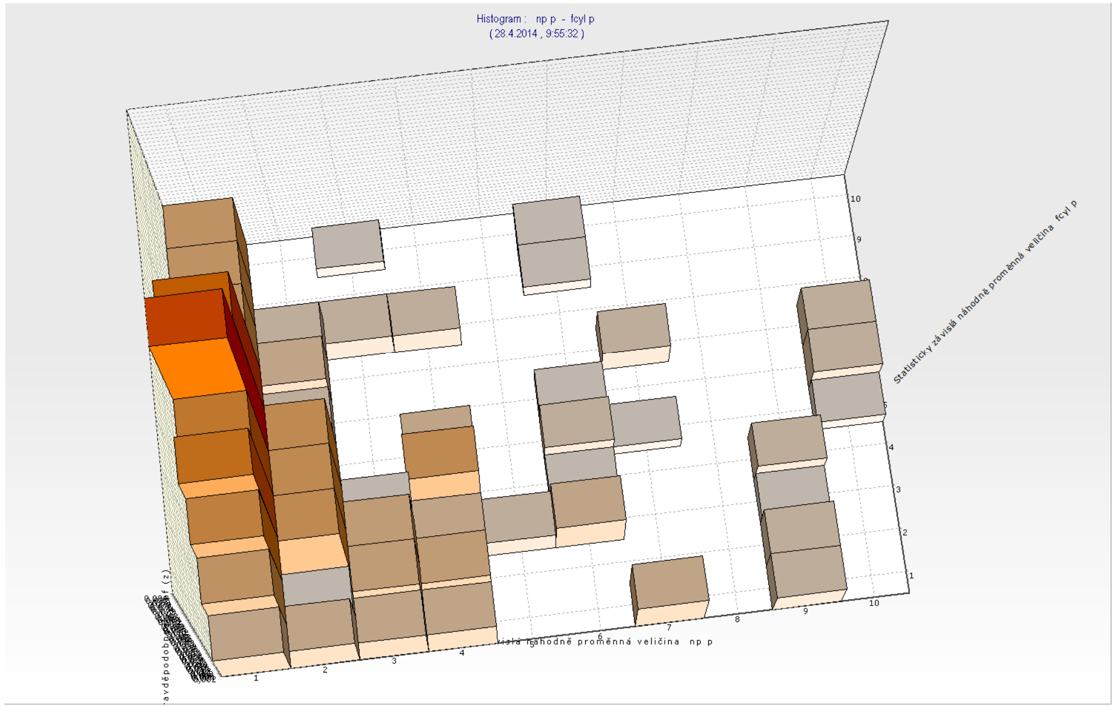


cube vs. cylinder compressive strength
the correlation 99.8% to 100.0%

Statistical Analysis of Dependence

Statistical analysis of concrete in the frame structure:

- A study based on strength properties of concrete depending on the floor of sample was also performed.
- The study pointed to a slight statistical correlation between the compressive strength of concrete and the ground floor, where the drill core sample was executed.
- The floor decreases with increasing compressive strength of concrete.



compressive strength of concrete vs.
floor in the building
the correlation -21.1% to -25.8%.