Moscow, Russia
14-17 April, 2020

Using the Data Scientist Experience in Training

Vladimir
Tupov
National Research University “MPEI”

Vladimir B. Tupov
Thermal Power Plants Department
National Research University “MPEI”
Moscow, Russia Federation,
TupovVB@mpei.ru

Using the Data Scientist Experience in Training

"Data is the new oil"

Clive Humby (2006)

"The world's most valuable resource is no longe[redacted], but data"

"Econo[redacted] (2017)

Using the Data Scientist Experience in Training

The possibility of using the experience of a Data scientist in the educational process using the example of acoustic calculations. The category of big data includes data streams is over 100 GB per day.

In acoustic calculations, comparable flows are from a large industrial conglomerate. In fact, the acoustic calculation in accordance with is performed for eight octave bands with geometric mean frequencies of 63, 125, 250, 500, 1000, 2000, [illegible]nd 8000 Hz and the sound level in dBA for each point.

The design points with a grid of 10 m for the surrounding area of 3000x3000 m near the heat power plant (TPP) will be 90,000, and the calculated values for one time point are 810000.

Using the Data Scientist Experience in Training

Each of the values will depend on several factors n, so we get an array of data

$$M=810000^n$$

These factors include *n*:

Number of working equipment;

Type of equipment;

Operating mode;

Equipment layout;

The height of the noise source above the ground;

Distance from the source of noise to residential buildings;

The presence of natural and artificial obstacles to the propagation of noise;

The orientation of the noise source in relation to the residential area;

The height of the calculated point above the ground;

Climatic factors and so on.

Thus, the amount is n> 10.

Using the Data Scientist Experience in Training

Acoustic calculations are carried out in many training courses, namely, “Environmental Technologies”, “Energy Gas and Air Pipelines of Thermal Power Plants”, etc. An acoustic calculation takes place in bachelor's and master's works.

It does not cover the entire huge volume and variety of data types. Therefore, the article does not claim to be absolute completeness and comprehensiveness. It is shown that the experience of a data scientist can be useful due to its versatility for training.

In fact, the data scientist category specialist s our main tasks:

1. Conversion of the source data into a form suita

2. Actual data analysis;

3. Interpretation of data;

4. Application of data to practice.

Using the Data Scientist Experience in Training

Conversion of the Source data into a Form Suitable for Analysis

Conversion of the source data into a form suitable for analysis is very important. The initial data for acoustic calculations are the values of sound power levels from hundreds of TPP sources, which are located both on the territory of the station and inside the TPP buildings.

Initial data can be provided by the Customer, manufacturers, but also obtained from the Internet. The last path becomes more important every year.

Characteristics of noise sources are often presente[illegible]nd pressure levels. Therefore, it is necessary to recalculate the sound pressure levels to the sound power levels from these sources by known methods.

To determine the radiated sound power from the air intakes of the blower fans, GTU compressors, it is necessary to take into account the decrease in the air duct, and from the cuts of the mouths of the chimneys, the decrease in the gas duct. When solving educational tasks, the most appropriate is to perform acoustic calculations according to established methods and the noise mapping.

The programming framework for programming distributed computing within the MapReduce paradigm is Hadoop MapReduce. For acoustic calculations can be used foreign software products, for example, SoundPLAN (USA), Predictor (Holland-Germany) and Russian software products – ARM-Acustica (LLC «TechnoProm»), Ecolog-Shum (the «Integral» company), etc

Using the Data Scientist Experience in Training

Actual Data Analysis

The software product allows you to obtain sound pressure levels from 63 to 8000 Hz and sound levels at any design point in a tabular form. An additional bonus is the presentation of the results of acoustic calculations in 3D, which allows you to determine sound levels by the height of buildings, take into account shielding from various industrial and residential buildings, as well as from the terrain. You can take into account the influence of the living area and orientation of the buildings on the noise level.

Students prepare a mathematical model of the energy  erforming calculations for self-created models' students analyze:

- dependence of the noise level in the surrounding area on the power and type of equipment, their quantity;
- noise levels in various layouts of energy facilities;
- the influence of the operating modes of various equipment on the noise level in the environment;
- the effect of the location of residential areas at different distances from the energy facility on the noise level;
- the influence of the orientation of residential areas relative to the energy facility on the noise level;
- the impact of climatic factors on reducing noise from equipment;
- the influence of the height of industrial, residential buildings on the distribution of sound.

Thus, the number of tasks has no restrictions on variety and complexity.

Using the Data Scientist Experience in Training

Interpretation of Data

The visual presentation of the results of big data analysis is crucial for their interpretation. The construction of the isolines of sound levels allows them to be compared with acceptable values . Therefore, it is clear to see where there is an excess of sanitary standards, and where not.

The calculation results can be presented in ta  rm in dBA and in octave bands. But for even better analysis use the constructed isolines of levels on the ground.

In the figure with contours, you can easily fix areas that exceed sanitary standards, and also where there are no excesses.

Visualize the results of acoustic calculations in 3D and determine sound levels by height.

Using the Data Scientist Experience in Training

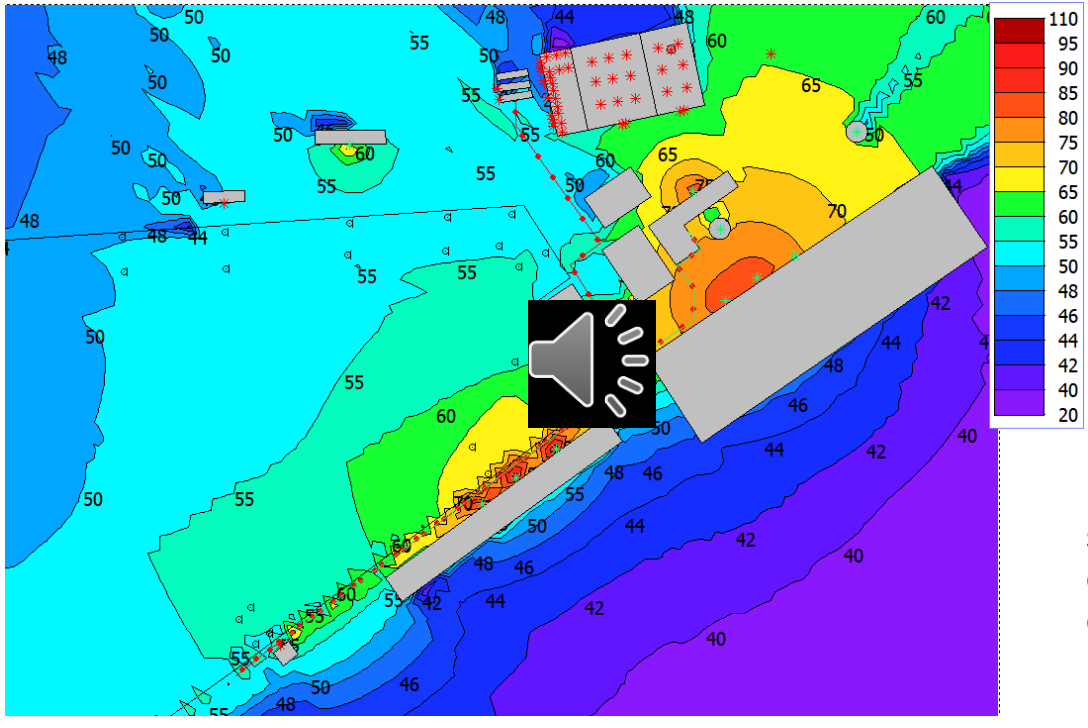
Application of Data to Practice

The obtained values have important applications. Calculations allow you to determine the required noise reduction from each type of equipment.

A set of measures to reduce noise is taken from the literature. The effectiveness of these measures is considered in the source data, on which repeated calculations are performed.,

It is possible to trace impact of various factors cha  d to choose an optimum way the noise attenuation. Examples are given below.

Using the Data Scientist Experience in Training



Isolines of sound level from “new” part of TPP: 1 —“old” part of TPP; 2— “new” part of TPP

Fig. 1 shows isolines of sound level from “old” part of TPP (sign “1” is “old” part of TPP, sign “2” is “new” part of TPP

Using the Data Scientist Experience in Training

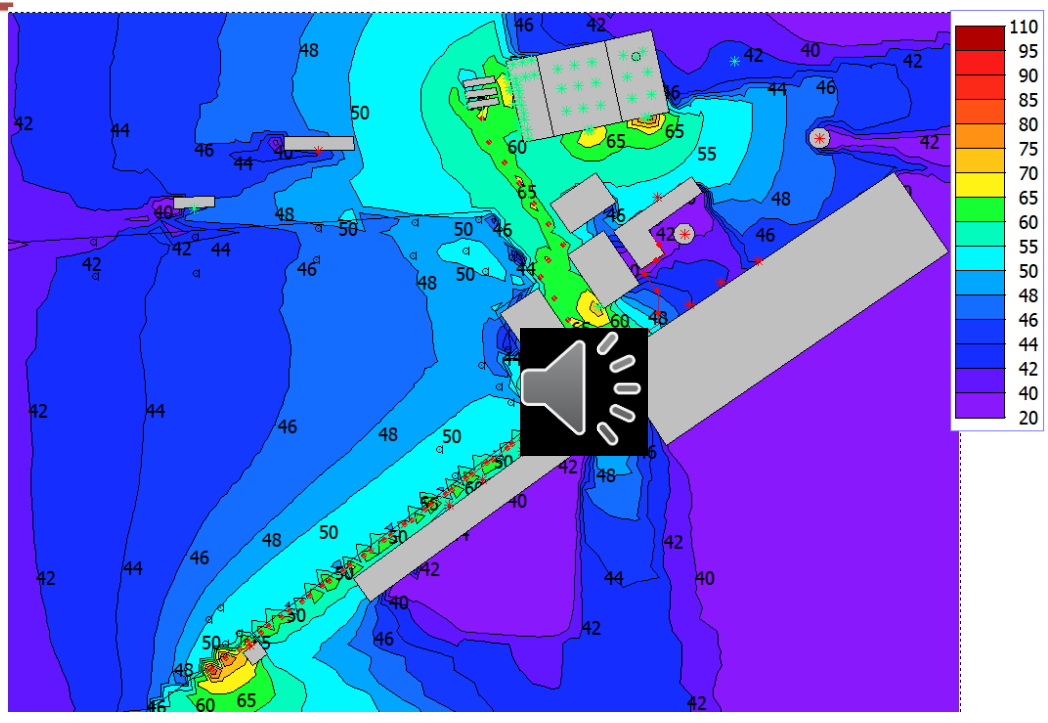
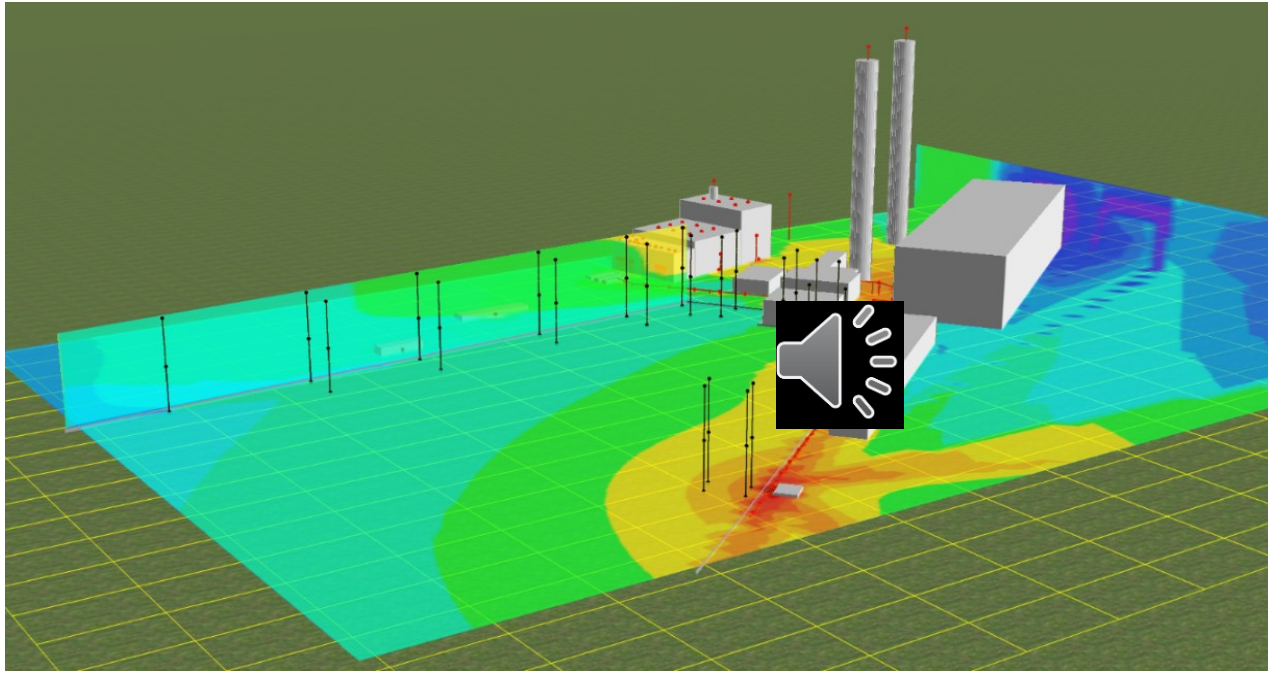


Fig. 2 shows isolines of sound level from “new” part of TPP (sign “1” is “old” part of TPP, sign “2” is “new” part of TPP).

Isolines of sound level from “new” part of TPP: 1 —“old” part of TPP; 2— “new” part of TPP

Article's Title



3D model and isolines of sound level in horizontal and vertical planes: 1 —"old" part of TPP; 2— "new" part of TPP

Fig. 3 shows the results of calculations of the sound level in the surrounding area in 3D. The results are presented in horizontal and vertical planes. As can be seen from the figure, sound levels vary greatly in the vertical plane, which is very important for practical measures to reduce noise.

Using the Data Scientist Experience in Training

CONCLUSION

The universality of the experience of a data scientist for training is shown on the example of performing acoustic calculations for large industrial conglomerates. This applies to all stages from the conversion of the source data into a form suitable for analysis, the actual analysis of the data to their interpretation and application of these data in practice. The importance of visua the obtained values in 3D for their interpretation is shown.

Moscow, Russia
14-17 April, 2020

Thank you for attention!

Speacker's contacts:

Vladimir

Tupov

National Research University "MPEI"

TupovVB@mpei.ru