The development of software to calculate the atmospheric pollution parameters and air pollution levels forecast

Azamat Kauazov [1,2], Nurlan Abayev [1,2], Shyngys Turashov[2] and Fariza Zhambalina[2]

[1] al-Farabi Kazakh National University, 71 al-Farabi ave., 050040, Almaty, Kazakhstan
[2] RSE "KAZHYDROMET", 11/1 Mangilik El ave., 010000, Nur-Sultan, Kazakhstan
kauazov_a@meteo.kz

Abstract. The article describes the development of applied software for calculating the atmospheric pollution index based on normalizing the observed impurities concentrations in the atmosphere and forecasting the level of pollution in Kazakhstan. Application software was developed for entering, calculating and automation fore-casting the level of pollution (P parameter) in 20 cities of the Republic of Kazakhstan. The automation of the P parameter calculations has significantly (several times) reduced the time spent on calculating the P parameter and the search for optimal pollution indicators in the city. The software is written in Borland Delphi 10 lite, the database is organized on the basis of MS Access. Database queries are executed in SQL using ADO technology. When the software runs, arithmetic calculations perform at the local level and output / input results are transported to MS Excel files. #CSOC1120

Keywords: database, programming, automation of data processing, software, application, air pollution, calculation, forecasting, P parameter.

1   Introduction

These days, the rising issue of air pollution in cities is undeniable. Due to the fact that the fossil fuel, burning of which releases toxic substances into the surface atmosphere, such as nitrogen dioxide (NO2), nitrogen oxide (NO), sulfur oxide (SO2) and others, is used as the main source of energy in cities. The hazardous products resulting from fuel combustion include nitrogen oxides and benzopyrene [1]. It is evident that contamination of the surface atmosphere not only negatively affects the environment in general, but also the state of human health [2].

The problem becomes worst in winter seasons, when emissions associated with heating increase that causes conditions named as Unfavorable Meteorological Conditions (UMC). In detail, UMC are the circumstances that prevent the dispersion of impurities and contribute to their accumulation in the atmosphere. Even though UMC (calm, inversion) has been observed with relatively small emissions for several days, the concentration could reach extreme values.

Measures on controlling and improving environmental conditions can be taken by using monitoring systems. The RSE "Kazhydromet" conducts air pollution monitoring in the main cities of Kazakhstan. It allows solving the following tasks related to air

basin quality management:
- control over compliance with state and international air quality standards;
- promptly receiving of objective information on the state of the air atmosphere;
- making decisions on the implementation of measures to reduce the level of air pollution based on the data obtained;
- modeling possible environmental conditions and predicting pollution levels;
- planning of urban development and transport systems;
- providing of information on-time to state authorities, environmental organizations and citizens about the air quality and the start of the warning systems about a sharp increase in the level of pollution [3].

However, the system of controlling and improving the state of the air basin of cities in Kazakhstan has become challenging due to the absence of a modern system for assessing pollution in cities and a nationwide system for predicting air pollution. Air pollution forecast is carried out only in few cities and on weekdays, since pollution forecast is carried out according to the data of laboratory observation points that do not work on weekends and holidays. Data from automatic observation posts were not used.

Recently, automation tools have become an integral part of work process, since they help to effectively use the working time, quickly access to the necessary information and storage, and they also provides a convenient user experience [4].

In addition, computer technologies and tools are of great importance in modern management and monitoring, ensuring the promptness of fixing, collecting, processing, searching and transmitting information, the reliability of its storage, remote access, providing information at the right time, on the right medium and in the right form as well as taking into account psychological and ergonomic requirements [5]. To assess and predict the level of pollution, it was necessary to solve several problems:
1) calculating the average concentration for different periods using the formula for calculating q.
2) identifying the maximum values exceeding the threshold q and MPC
3) dividing the concentrations by the MPC value and the operations of summation for the calculation by the formula Pnorm.

A conducted market research revealed that there is an absence of software products that could address similar problems. The software products available on the market are oriented towards standard calculations, and do not allow for research calculations, or are not oriented towards a system of operational monitoring and forecasting. For example, a well-known software in the field of environmental protection named as "Integral" - the Russian leader in software development, is focused only on calculating emissions according to approved methods for individual sources and does not allow its use in operational forecasting.

Therefore, the goal of this work was to develop Windows compatible application software (Application) for employees of the National Hydrometeorological Service of the RSE "Kazhydromet" to automate the calculating process of atmospheric pollution parameters. It was done by creating an application that allows data accumulation, data sampling upon request and processing followed by the sending of the results to users for scientific research and operational forecasts of the level of air pollution in the Republic of Kazakhstan.

2 Methods and data

To assess the pollution level in cities, the team used separate approaches such as assessing AQI pollution and forecasting pollution based on the parameter P. However, their usage had a number of drawbacks. For example, standards for AQI were set for only six substances, and most air pollutants do not have an associated AQI and require the development of limits for non-common impurities. To analyze the causes and consequences of atmospheric air pollution, it is necessary to use data on real concentrations of pollutants. In addition, it is difficult to apply the AQI and compare it with the officially adopted system for assessing the harmfulness of emissions based on the MPC.

According to the recommendation of the MGO the standard parameter P_{stand} is used as a characteristic of the background air pollution to systematize and assess the level of air pollution in the city as a whole.

Where n is the total number of observations of the concentration of impurities in the city during one day at all stationary points, m is the number of observations during the same day with concentrations q exceeding the average seasonal value $q_{mid-seas.}$ more than 1.5 times ($q > 1.5\ q_{mid-seas.}$).

The P_{stand} parameter allows one to estimate the level of pollution, and a forecasting system was created on its basis, but the P_{stand} parameter is calculated relative to the emission level of the previous year and therefore does not allow comparing different cities in terms of pollution level and assessing the dynamics of pollution.

Therefore, we made an attempt to calculate the parameter P normalized with respect to the MPC. The authors have developed a normalized pollution indicator in the form of the sums of the concentration of impurities in fractions of the MPC-$P_{norm.}$

To develop a forecasting system and identify the best parameter, it was necessary to carry out a large number of calculations to detect the statistical relationship be- tween meteorological parameters and atmospheric pollution parameters.

At the same time, it was decided not to be limited to the calculations of the parameter P_{stand} and P_{norm}, but to calculate other indicators to study the presence of correlations between the weather and pollution parameters. It was necessary to carry out a large number of calculations to develop predictive schemes. There was a need to calculate several pollution parameters for more than 120 observation posts in 20 cities of Kazakhstan, for several impurities in each city, on a daily basis for a period of several years. Meanwhile, the calculation of the P_{stand} parameter includes the calculation of some intermediate values and is quite laborious. It was not advisable to calculate such a number manually, even with the use of macros in MS Excel tables, since the test tables with macros were cumbersome and inconvenient to use. In addition to mass calculations for research work, there was a necessity to use the same algorithms for one-time calculations by the forecasting units of Kazhydromet for daily pollution forecasts in the long run. Thus, an urgent need arose to automate the calculations of pollution parameters and solve the following problems:
- improving the quality and productivity of work on the calculation of the parameter P and its forecast;
- creation of a unified database for the centralized storage of information on air pollution in the cities of Kazakhstan;

- providing of operational calculations of atmospheric pollution parameters to ensure the forecast of UMC with a high reliability degree;
- comprehensive calculation of various pollution parameters for research work to identify the best indicators;
- release of the user from all routine and time-consuming operations to calculate the pollution parameter P.

The research object is the process of calculating the parameters of atmospheric pollution-P, which takes into account the emissions of harmful substances into the atmosphere in the city as a whole. The main data used is presented in the form of air pollution observation data from the network of laboratory and automatic observation points of Kazhydromet (National weather service). The data on pollution is presented in the standard Air Pollution Observation Tables (APOT-1) format. Such tables record observation data for each observation post during a month for several impurities. Also, when calculating related to the forecast, the program operates with the values of air temperature and wind speed at altitudes. Calculations represent the simplest calculations, as well as interaction with MS Excel tables.

For the given task, it is needed to perform simple operations and calculations. The mathematical tool is presented as simple arithmetic operations, as well as the calculation of the number of maximum values for the given thresholds. The search for extreme values is carried out by the method of sequential comparison of each concentration with the specified thresholds (q, MPC). When an element is found with a value greater than the threshold, the corresponding value in the counter variable and variable is incremented. At the end of the search, a report with a calculation result and the number of elements corresponding to the search requests will be presented in the MS Excel file.

In general, when developing an information system model, methods such as abstractlogical, systemic and comparative analysis, methods of analysis and optimization of business processes were used.

To create such an application that will satisfy the task at hand and ensure the functioning of the mathematical apparatus, standard components were used:

1) One component for visual presentation of data in tabular form;
2) Fields, calendar for entering the parameters necessary for the calculation and additionally for visual confirmation of the actions performed by the user;
3) A set of checkboxes and radio buttons for selecting calculation parameters, information about that is stored in the database;
4) A dialog box that provides a selection of files with the appropriate buttons to call them;
5) Components for connecting to the database;
6) Component for queries to the database (DB) based on ADO technology;
7) Built-in browser for viewing help files.

Other components of the program should provide data query and output, data editing, data import from MS Excel tables into the database, fill in and copy the values of cells in MS Excel tables.

3   Results

3.1   Application description

When starting the application, the user sees the splash screen and the main form (Fig. 1), which contains tabs and menu items. The developed software contains four modules: Entering and viewing data, Calculation of UMC, Forecast of UMC and Bulletin.

1.   The Data Entry and Viewing Module (see Fig. 1), contains reference and informational data on pollution (concentrations) and the ability to display information about
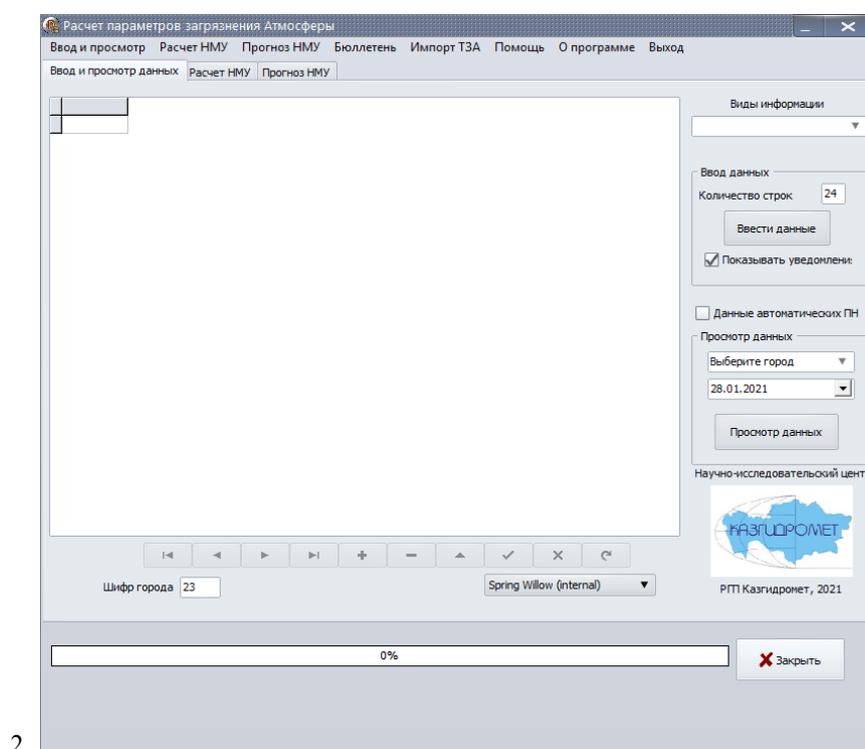
2.

Fig. 1. Main window of the Application, View and data entry tab.

the concentration by city for a date on request in a table for viewing. Data loading is carried out only from external files. To optimize work, data loading is carried out from standard APOT-1 reporting forms. To load information into the database, the user must enter the number of lines required for import, select the external Excel file of the APOT-1 and click on the "enter data" button. Data on the entered concentrations of pollutants in the city, by posts, at various times for a certain day can be viewed in the table.

The user can not only add data to the database, but also delete it. Removal is done through the standard navigation bar. The user can click on the "-" button to delete data, as a result the selected record is deleted from the database. The user cannot change the

data.

1. The user selects the search and calculation criteria (city, date, impurities), as well as the calculation period (day, month, year) and clicks the corresponding Calculation button for parameter P. Next, the software accesses the database with a request for pollutant concentrations for a certain period for certain criteria. Next, the calculation of the sums of the concentrations of impurities in fractions of the MPC, and dividing by the total number of observations is made. The algorithm of the program's operation consists in a SQL query for information about the level of pollution for a certain peri- od of time, for this the program dynamically generates a code sequence which, by means of a dynamically generated SQL query, forms criteria (value q) which are then sequentially compared for coincidence with a data set for a period selected by the user.

2. Module Calculation of UMC (see Fig.2) contains a list of cities and control elements for selecting parameters and calculation period. In case of coincidence (exceeding) of the threshold values, the corresponding value of the behavior anomaly index is formed, entered into memory / file, and the procedure is recursively repeated until the conditions of the cycle are met. The result is displayed either in a pop-up message for a day, or in an Excel file for a month and a year.
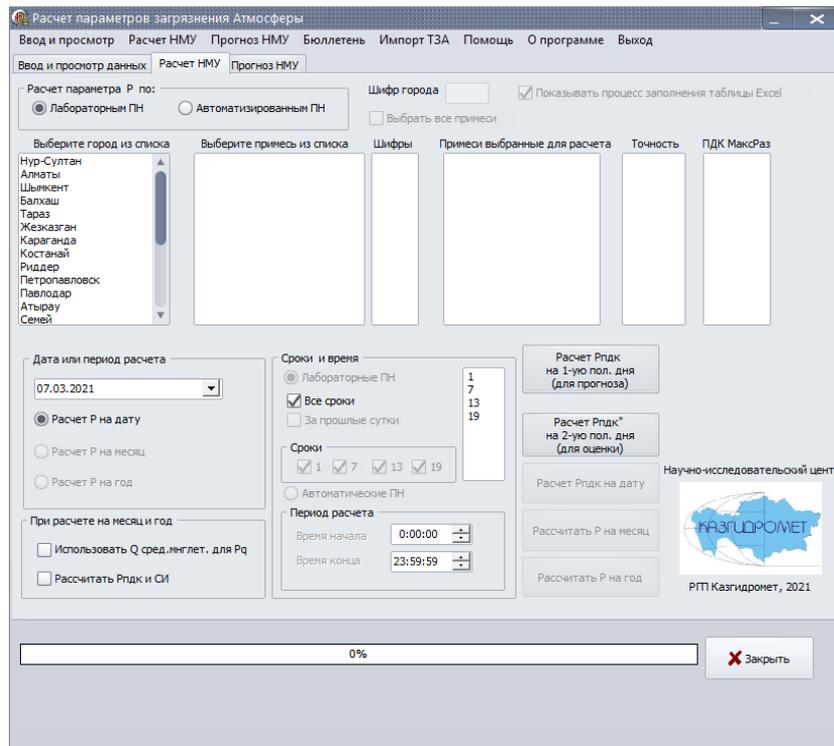

Fig. 2. Main window of the Application, tab Calculation of UMC.

3. The Design Bulletin module is similar to the NMU tab but in a simplified ver- sion. The user operating procedure and the algorithm calculation are similar to the operating procedure and the algorithm calculation of UMC. The difference is that the final result is compared with critical values and is automatically compared to one of three different pollution groups and the output is transported to a special template file, where the daily newsletter template is partially filled in.

4. The forecast module is focused on calculating the predicted P value. The user se- lects the city and date (see Fig.3). Then he enters the actual data on the temperature and wind speed near the ground and at heights (925 mb-800m, 850mb-1500m) obtained from the meteorological station and upper-air sounding data at 00 UTC. It also intro- duces predictive data on the temperature and wind speed at the ground and at heights (925 mb-800m, 850mb-1500m) obtained from WRF models for 12, 24, 36 and 48 hours. The P values for the current and previous day are also entered. Next, you need to press the button Calculate P.

Fig. 3. The main window of the Application, the Forecast tab

After that, the application automatically fills in the forecast table in Excel format, where all the calculation formulas have already been entered and reads the values from the cells with the calculation results. Further, the forecast results are copied to another table named as "Forecast justification journal", which is partially filled in automatically. The

second part of the journal is filled in and the final forecast is formed by the forecaster based on his vision of the synoptic situation development.

For the convenient user experience, there is a help system. The main menu also has an "import data" option. Monthly data import can be done through the menu item called Import APOT-1.

To select Excel files when importing data from APOT-1 tables, regular Open Dialogs are used to provide the user with the ability to visually select the file name and path.

To duplicate functions and create additional convenience in working with the software, the Main menu component was implemented with the corresponding submenus and button captions. The ADO Connection component was used to connect to the database and ADO Query and Data Source to query to the database.

3.2   Database description

MS Access DBMS was used for data storage. To build an infological data model, at the first stage, the entities of their relationships and attributes are determined. During the analysis of the subject area the main functions of the system such as entering data on pollutants concentration, searching for data on pollution, and calculating a parame- ter with outputting the result to an external file were highlighted. The domain data was structured as a single entity. In this case, the entity is represented by pollution data. Entity attributes include concentration of various impurities, place (city and post), as well as date and time (year, month, day, period). Hence, the infological and datological model is quite simple: for each city, only one table is used, which presents data for the entire period and for all meteorological posts. There are no transitive and multivalued dependencies. The simplicity of the structure is associated with the need to increase processing request speed, as well as for an easier data perception.

Database queries are performed using SQL (Structured Query Language - a language of structured queries - a universal computer language to create, modify and manage data in relational databases). The database is represented by a set of tables with pollution data for the main cities of Kazakhstan and auxiliary tables of conditionally constant characteristics in the separate table. Each table has its own number that corresponds to a specific city. The structure of the table for the city is standard. There are fields with the date, separately broken down by components, a field with the number of the post and the type of post, the observation period or time, then in the fields with the name PN there are data on contamination for individual impurities, where P is not a variable designation of an impurity, and N is impurities' identification number. There is a separate dataset for each city. Following that, meteorological data on temperature, wind speed and direction, as well as the code of atmospheric conditions are represented. Storage and recording of data on impurities' concentration from laboratory and automatic observation points is carried out in different tables.

Excel summary tables are the output of the program. There are three types of summary tables:
•   Table with the pollution parameters calculation results for a month, a year and more.

• A table that is a partially formed daily environmental bulletin, where fields with an assessment of the level of pollution are automatically filled in and reference information about posts is indicated.

• The table that is known as forecast report automatically fields with predicted values of the P parameter according to two methods and data on temperature and wind for individual isobaric surfaces also the date and place of observation.

4   Discussions

While the algorithmic language Object Pascal was chosen to create the program in the visual programming environment Delphi, Delphi 10 Lite v.3 visual programming environment was used to further develop the program. Modified Alpha Controls components were used to improve the visual appearance. Our team decided to choose Delphi as it is a multi-purpose programming environment that allows users to develop a software product of any complexity for any area. The program can be rather successfully used by both professionals with an expert knowledge and graduates with basic fundamentals in programming. Delphi offers a fairly wide range of components for working with databases. The hierarchy of components for working with databases is organized in such a way that it practically does not matter which database the application uses. This feature enables Delphi to provide Rapid Application Developing (RAD). Such tools allow users quickly create a work program from ready-made components, avoiding a waste of massive effort on trifles. We assume that this Delphi feature has led to its widespread use in the development of automated enterprise management systems [6].

As a result, the aim of this work is to create an application that provides information processing and output. The application should provide work with database, information on pollution, perform mathematical calculations for calculating pollution parameters using one-dimensional data sets and provide the user with fast sequential calculations of several atmospheric air pollution indicators, both for a single period of time and for several years. It is also necessary to transport information output in MS Excel to facilitate further statistical analysis.

The degree of rationality of creating this program is assessed as high. Based on the initial conditions, it is necessary to perform a sequence of actions, such as:
- ensuring the storage of the average amount of data in the database;
- data request;
- conducting of simple mathematical calculations;
- output of the results.

When examining possible approaches, it was assumed that using open source software, based on Excel macros or using the Access DBMS tools would give an opportunity to create almost the same application that would be able to solve the given task. Although the calculation procedure becomes more complicated minimum effort would be spent on interface development. The main weakness of this approach is the lack of an interface when working with macros. Another drawback is the implementation complexity of the calculation algorithm for some parameters, or the need for skills in

working with the Access DBMS, which is not widely used in the Republic of Kazakhstan. Since the calculation should take place in an automated mode, with minimal user involvement, it was decided to create a separate application. Thus, the developed Windows application has a clearly defined purpose and place of use.

As indicated above, the function-oriented design approach was chosen as the methodology [7]. Using an object-oriented approach in programming, allows the user to reuse code with greater efficiency. Visual design tools for user interfaces allow you to speed up work on projects, facilitate code reuse. The application can be used for various purposes.

Firstly, the software can be used as a resource database that allows users to provide complete information on a specific city for an arbitrary period of time. Secondly, it can be helpful for scientific research purposes, or tasks where a large number of calculations of pollution parameters are required.

5   Conclusion

With the help of the functionally oriented design methodology an application software was created for entering, calculating and forecasting the level of pollution in 20 cities of the Republic of Kazakhstan. The total amount of program code is over 3,000 lines.

The developed software has no predecessors in the market and has a number of advantages such as:

1.   Simple and understandable components and terms that make it easier to learn and implement the program for potential users;

2.   A cost-effective and timesaving technical support. The software meets all the necessary requirements and provides a convenient and an effective application in practice. Also, the developed software made it possible to identify excess concentrations, excess frequency and their consistencies and contributed to the establishment of a closer relationship between atmospheric and pollution parameters resulting in a more accurate forecast.

3.   Useful calculations to predict atmospheric air pollution in the cities of Kazakhstan.

Overall, the automation of the P parameter calculations has significantly (several times) reduced the time spent on calculating the P parameter and the search for optimal pollution indicators in the city. The automation of processes of bulletins and forecasting journals formation has greatly simplified the work of chemists, laboratory assistants and forecasters. The software has been handed over to branches and successfully implemented to make pollution predictions. The program successfully works in Windows 2000 and higher operating systems by using standard components and with no additional expensive software installations need. It is planned to enhance the program and introduce automatic calculations and data loading in the long run.

References

1. Zlotin G.N., Griga A.D., Kulanov V.A., Griga S.A. The joint effect of nitrogen oxide and benzopyrene emissions during the operation of power-generating boilers at CHPPs // Bulletin of the Volgograd State Technical University. 2009. No. 7 (55). pp. 34-36.
2. Murtazina M.Sh., Panchenko M.A. Development of an information system for assessing the negative impact of energy enterprises on the whole by emissions of harmful substances into the atmosphere // Internet magazine "Science" Volume 8, No. 5 (2016) http://naukovedenie.ru/PDF/40TVN516.pdf , last accessed 2021/18/01.
3. Volosova Anastasia Sergeevna, Sverchkov Andrey Mikhailovich Software application for collecting and analyzing data on the ecological state of the atmosphere of the city of Moscow // Successes in chemistry and chemical technologies. 2017. No. 1 (182). URL: https://cyberleninka.ru/article/n/programmnoe-prilozhenie-sbora-i-analiza-dannyh-ob-ekologicheskom-sostoyanii-atmosfery-goroda-moskvy, last accessed 2021/18/01.
4. Koroleva, I. Yu. Automation of the process of development of the UMKD department of the university / I. Yu. Koroleva, DG Vlazneva. - Text: direct // Young scientist. - 2012. - No. 3 (38). - S. 92-95. - URL: https://moluch.ru/archive/38/4445/, last accessed 2021/18/01.
5. Fedotov Mikhail Yurievich, Elizarova Ekaterina Nikolaevna Development of an information system for accounting and sales of goods in agriculture in the Delphi environment
6. //International Scientific Review. 2017. No. 4 (35). URL: https://cyberleninka.ru/article/n/razrabotka-informatsionnoy-sistemy-dlya-ucheta-i- realizatsii-tovara-v-selskom-hozyaystve-v-srede-delphi, last accessed 2021/12/02.
7. Designing information systems: textbook. allowance / V.I. Grekul, G.N. Denischenko, N.L. Korovkin. - 2nd ed., Rev. - S .: Internet University of Information Technologies; BINOMIAL. Knowledge Laboratory, 2008 .-- 300 p.