

Modelling of non-point source pollution transport for the Charyn River Basin

Jalal K Jamalov*, Daniyar B Nurseitov, Kairat A Bostanbekov

National Open Research Laboratory of Information and Space Technologies, Kazakh National Research Technical University named after K.I. Satpayev., Satpayev Str. 22a, Almaty, Kazakhstan

*Corresponding author: gg.mazei@gmail.com

Received 26 December 2016, www.cmnt.lv

Abstract

The results of pollution transport simulation for the Charyn river (the Republic of Kazakhstan, Central Asia) obtained using software package BASINS 4.1 are shown in this article [1]. Modules created in the process of the study as well as the method of adaptation of the model of pollution transport are described. The calculations include the modeling of the hydrology of the river basin and the calculation of the concentration of non-point sources of pollution. The comparison with the data of natural hydrological observation post.

Keywords

Simulation
BASINS
Watershed Delineation
HSPF
pollution transport in water
BOD
nitrate
dissolved oxygen

1 Introduction

The Charyn is the river in Almaty region of the Republic Kazakhstan. It is located in the Charyn gorge where the Charyn canyon belongs. It is the largest left-bank tributary of the Ili river. Its headwaters are located above the climatic snowline, on the southern slope of the Ketmen range. The river length is equal to 428 km, catchment area - 9.0 thousand square kilometres. The main tributaries are Karkara (the right one) and Temirlik (left). Currently, monitoring of water flow is being carried out by the only remaining station of RSE "Kazhydromet" on the Charyn river located in Sarytogay tract. According to the data of this station the average annual water discharge makes up 37.7 m³/s.

2 Theory and calculations

The region under study is the basin of the Charyn river shown in Figure 1. To implement calculations multifunctional system of environmental analysis BASINS 4.1 was used. This system was developed by the Environment Protection Agency for the purposes of setting standards of the maximum total daily load (TMDLs) for waterbodies with impaired water quality and to allow local and state agencies to conduct analysis of watersheds. To simulate the transfer of pollution the main input data were determined, and the algorithm for calculations was developed. In Figure 2 a graphical model describing the developed algorithm is given.

Input parameters for BASINS 4.1 software are the following ones:

- 1) Digital Elevation Model (DEM);
- 2) River Net;
- 3) Land Use data;
- 4) State Soil data;
- 5) Meteorological data.

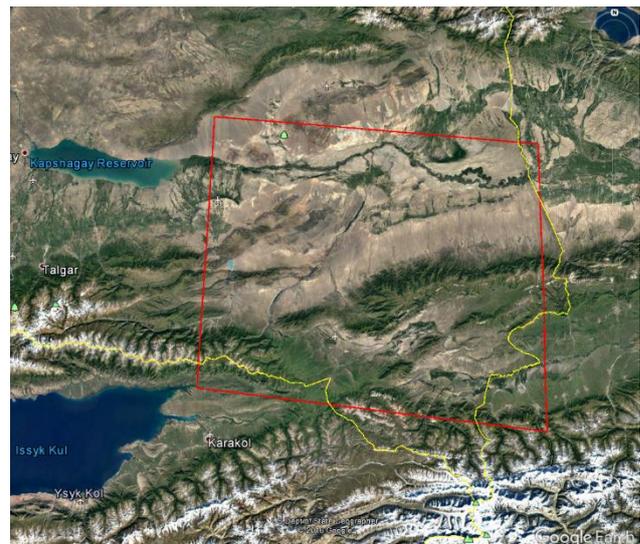


FIGURE 1 The region under study

The digital elevation model (DEM) is the information on terrain elevations in the area under study presented in GeoTIFF format. This format is an input one to be used in the Watershed Delineation tool (WDT) [2] intended for determination of the boundaries of the basin and the construction of watershed lines. The major part of DEMs is generated by the processes involving land remote sensing technology, the search and analysis of online resources that may provide this information were conducted. As a result, the DEM of the Global Explorer Data portal was used [3]. To construct the lines of watersheds and to determine the basin boundaries in addition to elevation data (DEM) the data of river flows location (River Net) were used. The USGS HydroSHEDS has developed files of streams and river networks based on digital elevation models obtained by NASA [4] (in the course of SRTM).

The data are given in vector format and contain the information on the location of river flows of the entire Asian region.

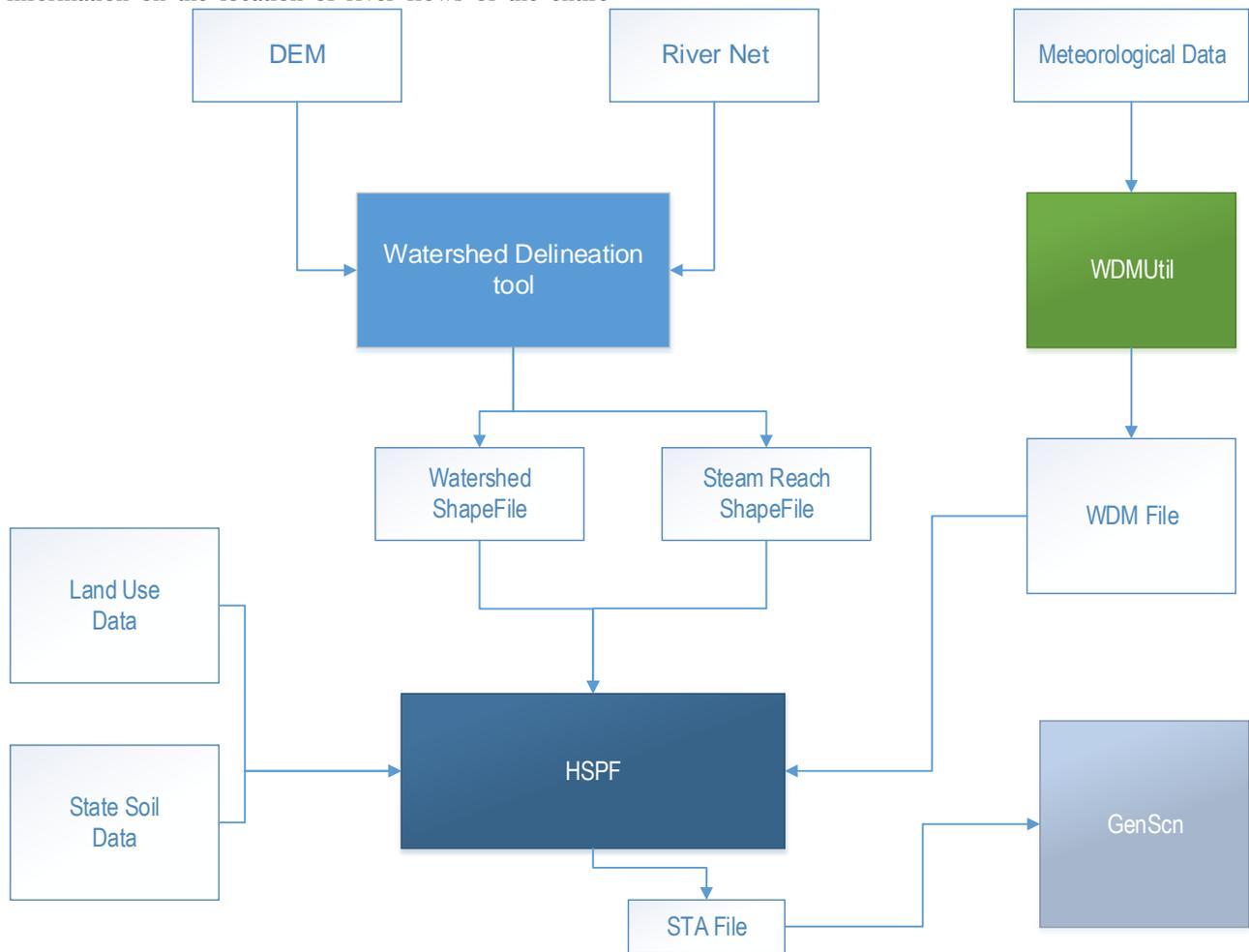


FIGURE 2 Flow chart of calculations

The results of the WDT usage are:

- 1) basin boundary lines;
- 2) sub-basin lines;
- 3) river network line file.

The files were visualized using Google Earth, where: 1) the boundary of the river Charyn basin (translucent blue color); 2) river Charyn (blue marker); 3) sub-basins of the river Charyn (red marker), see Figure 3.

For the purposes of hydrological processes simulation HSPF model was used. This model requires the following input data:

- 1) sub-basin lines;
- 2) the river network lines;
- 3) land-use maps;
- 4) soil maps;
- 5) meteorological data.

The application of the data of land cover in BASINS allows evaluating the categories of land use in the watershed. To minimize the errors of the HSPF model results land cover data resolution should be as high as possible, especially for small catchment basins. Land use maps Global Land Cover 2000 were downloaded in TIFF format from the website of the Joint Research Centre of the European Commission (Joint Research Centre) [5]. Soil data were obtained from the FAO GeoNetwork website [6].

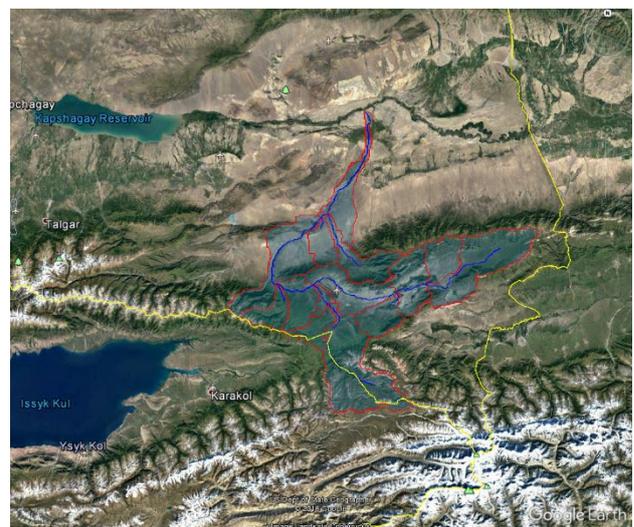


FIGURE 3 Image by Google Earth with superimposed data of calculations

In HSPF model land use and soil maps are used for determining the area and hydrological parameters of each category of soil and land resources. Application of the tool Spatial Analyst allows calculating cross areas between tabular data of land cover and sets of soil cover by the method of superimposition.

3 Meteorological data processing

As for hydrological processes, these vary over time and depend on the changes in the state of environment, while for evaluation of pollution of scattered sources the succession of hourly precipitation, evaporation, temperature and other meteorological data are required.

Initially, meteorological data were downloaded from the website of the National Climatic Data Center (NOAA's National Climate Data Center) [7]. Then, meteorological data of ALMATY station were used (marked with a yellow square), Figure 4.



FIGURE 4 Meteorological station ALMATY

The obtained data are stored in a text file (txt), Figure 5.

1	Identification	WIND	TEMP	DENPT	PRECIP	PRECIP	PR
2	USAF NCDC Date	HrMn I Type QCP	Dic Q I Spd Q Temp Q	Denpt Q Pr Amt I Q Pr	Amt I Q Pr	Amt I Q Pr	
3	368720 99999 20060101 0000 4 FM-15	230 1 N 8.0 1 8.0 1	-9.0 1 99 999.9 9 9 99 999.9 9 9 99				
4	368720 99999 20060101 0030 4 FM-15	230 1 N 6.0 1 6.0 1	-8.0 1 99 999.9 9 9 99 999.9 9 9 99				
5	368720 99999 20060101 0100 4 FM-15	220 1 N 6.0 1 8.0 1	-8.0 1 99 999.9 9 9 99 999.9 9 9 99				
6	368720 99999 20060101 0130 4 FM-15	210 1 N 6.0 1 8.0 1	-6.0 1 99 999.9 9 9 99 999.9 9 9 99				
7	368720 99999 20060101 0200 4 FM-15	230 1 N 3.0 1 8.0 1	-6.0 1 99 999.9 9 9 99 999.9 9 9 99				
8	368720 99999 20060101 0230 4 FM-15	190 1 N 2.0 1 6.0 1	-4.0 1 99 999.9 9 9 99 999.9 9 9 99				
9	368720 99999 20060101 0300 4 FM-15	160 1 N 2.0 1 6.0 1	-4.0 1 99 999.9 9 9 99 999.9 9 9 99				
10	368720 99999 20060101 0330 4 FM-15	120 1 N 3.0 1 6.0 1	-3.0 1 99 999.9 9 9 99 999.9 9 9 99				
11	368720 99999 20060101 0400 4 FM-15	110 1 N 3.0 1 8.0 1	-3.0 1 99 999.9 9 9 99 999.9 9 9 99				
12	368720 99999 20060101 0430 4 FM-15	999 9 V 1.0 1 8.0 1	-2.0 1 99 999.9 9 9 99 999.9 9 9 99				
13	368720 99999 20060101 0500 4 FM-15	180 1 V 2.0 1 7.0 1	-2.0 1 99 999.9 9 9 99 999.9 9 9 99				
14	368720 99999 20060101 0530 4 FM-15	999 9 V 1.0 1 7.0 1	0.0 1 99 999.9 9 9 99 999.9 9 9 99				
15	368720 99999 20060101 0600 4 FM-15	999 9 C 0.0 1 8.0 1	-1.0 1 99 999.9 9 9 99 999.9 9 9 99				
16	368720 99999 20060101 0630 4 FM-15	999 9 V 1.0 1 9.0 1	1.0 1 99 999.9 9 9 99 999.9 9 9 99				
17	368720 99999 20060101 0700 4 FM-15	080 1 V 2.0 1 11.0 1	-1.0 1 99 999.9 9 9 99 999.9 9 9 99				
18	368720 99999 20060101 0730 4 FM-15	999 9 C 0.0 1 11.0 1	-1.0 1 99 999.9 9 9 99 999.9 9 9 99				
19	368720 99999 20060101 0800 4 FM-15	350 1 N 2.0 1 10.0 1	1.0 1 99 999.9 9 9 99 999.9 9 9 99				
20	368720 99999 20060101 0830 4 FM-15	350 1 N 3.0 1 10.0 1	0.0 1 99 999.9 9 9 99 999.9 9 9 99				
21	368720 99999 20060101 0900 4 FM-15	350 1 V 2.0 1 8.0 1	1.0 1 99 999.9 9 9 99 999.9 9 9 99				

FIGURE 5 Meteorological data of the NOAA's National Climate Data Center

As it is seen in figure 5, the value of PRECIP = 999.9, that implies the lack of the data on precipitation, as well as on evaporation. However, these two parameters are very essential ones for calculations of water discharge and simulation of hydrological processes. Therefore, meteorological data to be used in the BASINS system were downloaded from the database of reanalysis ERA Interim, created by the European Centre for Medium-Range Weather Forecasts (ECMWF). The data are presented in GRIB format for the period from 1 January 1980 to 1 August 2016. The IDV (Interactive Data Visualization) software allowed exporting meteorological parameters concerning the point of 43.2 latitude and 79.0 longitude into CSV text format intended for presenting tabular data. The tabular data are stored in the form of time series, where the first column is the date and time, and each subsequent column is the value of the corresponding meteorological parameter. Since the model requires hourly time series, while the downloaded data were presented with a 6-hour interval, the values were interpolated. For this purpose, a script in Python for reading the data from the CSV file and conducting linear interpolation was written. Then the obtained values with hourly intervals and time shifting for +6 hours, as the initial data correspond to UTM, were stored in a new CSV. The python code listing is shown below.

```
import os
import sys
import math
import numpy as np
from scipy.interpolate import interp1d
import matplotlib.pyplot as plt
from matplotlib.colors import LogNorm, ListedColormap, BoundaryNorm
import datetime

try:
    infile = sys.argv[1]; outfile = sys.argv[2];
except:
    print "Usage:",sys.argv[0], "infile outfile"; sys.exit(1)

# read infile
fin = open(infile,'r')
header = fin.readline()
lines = fin.read().split('WrWn')
values = []
for line in lines:
    values.append(line.split(';')[1])

y = values
x = np.linspace(0,len(y)*6-6,len(y))

f = interp1d(x,y)
#f2 = interp1d(x,y,kind='cubic')

xnew = np.linspace(0,len(y)*6-6,len(y)*6-5)
ynew = f(xnew)
#ynew2 = f2(xnew)

#plt.plot(x, y, 'o', xnew, ynew, '--', xnew, ynew2, '-')
#plt.show()

fout = open(outfile,'w')
fout.writelines(header)

Time = datetime.datetime(1979,1,1,9,0,0)
for line in ynew:
    fout.writelines(Time.strftime('%Y-%m-%d %H:%M:%S')+';'+ '{0:10f}'.format(line)+'WrWn')
    Time += datetime.timedelta(hours=1)

fin.close()
fout.close()
```

For structuring time series of meteorological data WDM format was used [reference]. The characteristic features of

WDM files are the following ones:

- 1) these files are supported in HSPF model;
- 2) allow storing a large number of time series;
- 3) ensure reading, movement of the data stored in the file. To import the data from test files into WDM format a script for utilities WDMUtil was written.

```
(ATCScript "WDMUtil Export Format - Hourly Values"
  (LineEnd LF)
  (NextLine)
  (ColumnFormat Fixed
    1-4:Year
    6-7:Month
    9-10:Day
    12-13:Hour
    15-16:Minute
    21-36:Value
  )
  (Attribute Scenario "OBSERVED")
  (Attribute Location "Charyn")
  (Attribute Constituent "ATEM")
  (Attribute Description "Temperature")
  (Attribute LATDEG "43.2")
  (Attribute LNGDEG "79")
  (Test (And (IsNumeric Value)
    (IsNumeric Year)
  )
  (While (Not EOF)
    (Date Year
      Month
      Day
      Hour
      Minute)
    (Value Value)
    (NextLine)
  )
  (Fill H 1 -999 -999 -999)
)
```

```
(IsNumeric Month)
(IsNumeric Day)
(IsNumeric Hour)
(IsNumeric Minute)
(> Year 1700)
(< Month 13)
(> Day "0")
(< Day 32)
(< Hour 26)))
(While (Not EOF)
  (Date Year
    Month
    Day
    Hour
    Minute)
  (Value Value)
  (NextLine)
)
(Fill H 1 -999 -999 -999)
)
```

As a result, a WDM-format file with the required meteorological parameters, in the required units with hourly intervals was obtained, see Figure 6.

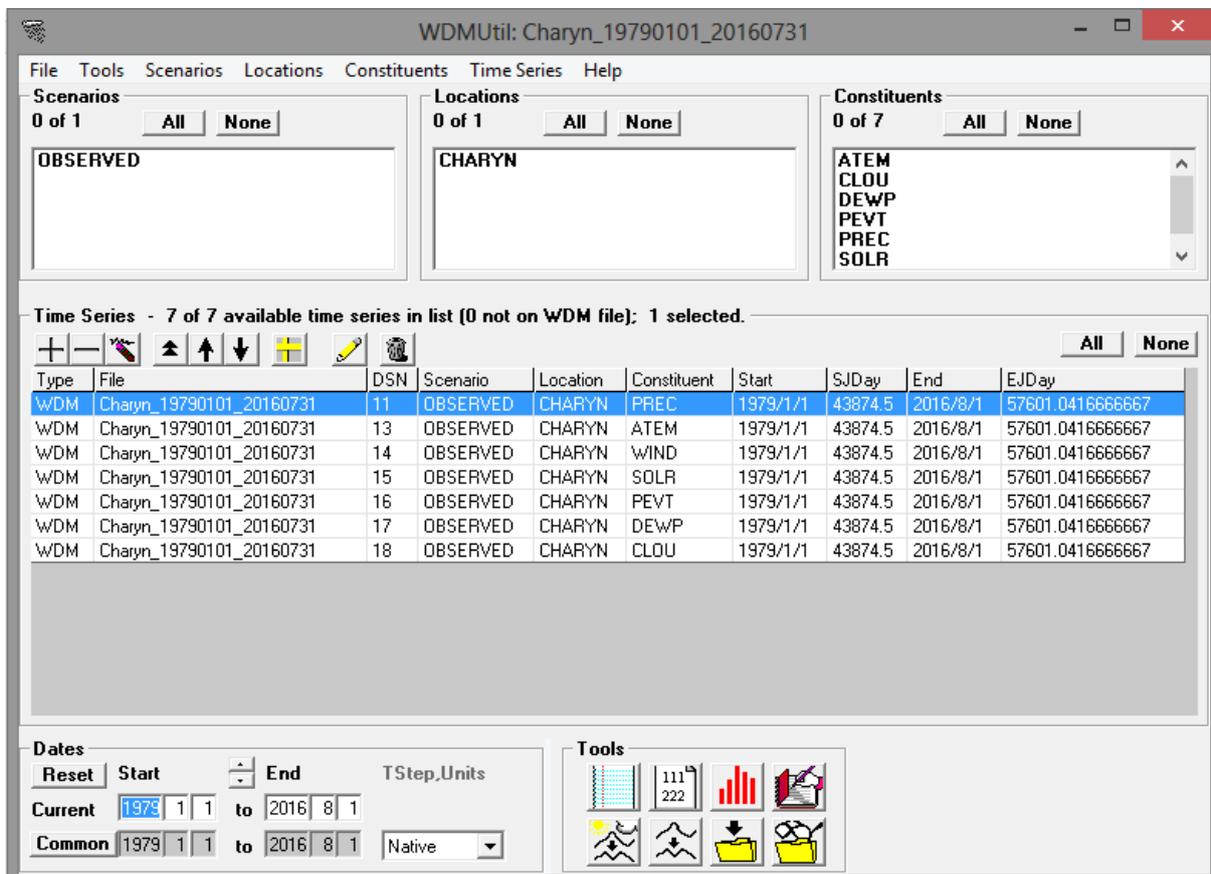


FIGURE 6 Meteorological parameters in WDM format opened by means of WDMUtil

As shown in Figure 6 in the “Constituent” column indicated all meteorological parameters used to calculate:

TABLE 1 Description of meteorological parameters

Data set	Description Parameter
PREC	hourly precipitation
CLOU	hourly cloud cover
ATEM	hourly temperature
WIND	hourly windspeed
SOLR	hourly solar radiation
PEVT	hourly potential evapotranspiration
DEWP	hourly dewpoint temperature

As a result of calculations in HSPF model for the River Charyn were obtained 15 river reaches of each of these specific parameters, see Figure 7.

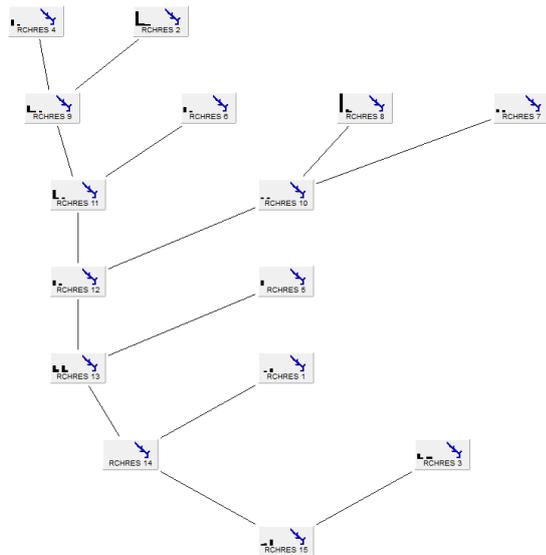


FIGURE 7 The Charyn river reaches in HSPF model

Reaches correspond to the main flow of the river and its lateral tributaries, as shown in Figure 8.

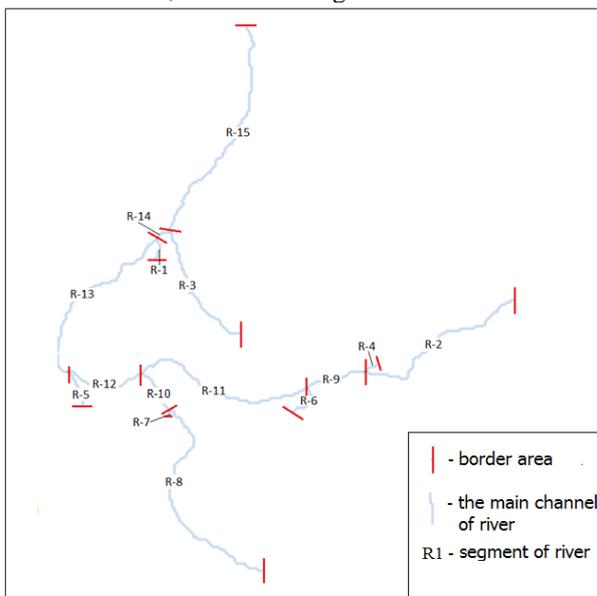


FIGURE 8 Schematic representation of the results of HSPF model

As seen in Figure 8, river reach R-15 is the final one. It is the area where the hydrological post - Sarytogay tract - monitoring the river water flow is situated. This fact allowed comparing the simulation results with the data of field observations.

4 Results

The following indicators were found for the following non-point sources of pollution:

- 1) BOD (Biochemical oxygen demand);
- 2) Dissolved oxygen (Dissolved oxygen);
- 3) Nitrate nitrogen (NO3);
- 4) Ammonium Nitrogen (Total Ammonia) (TAM).

4.1 BIOCHEMICAL OXYGEN DEMAND

Biochemical oxygen demand is the amount of oxygen consumed for aerobic biochemical oxidation by the action of microorganisms and decomposition of unstable organic compounds contained in the tested water. According to the observation data for the year 2006 obtained at the station located in Sarytogay tract, monthly average BOD concentration makes up 0.81 mg/L.

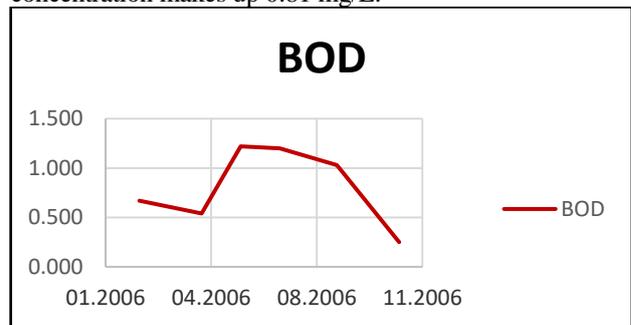


FIGURE 9 Changes in the concentrations of BOD

Average annual BOD concentration at reach R-15 obtained as a result of calculations equaled to ~ 0.72 mg/L.

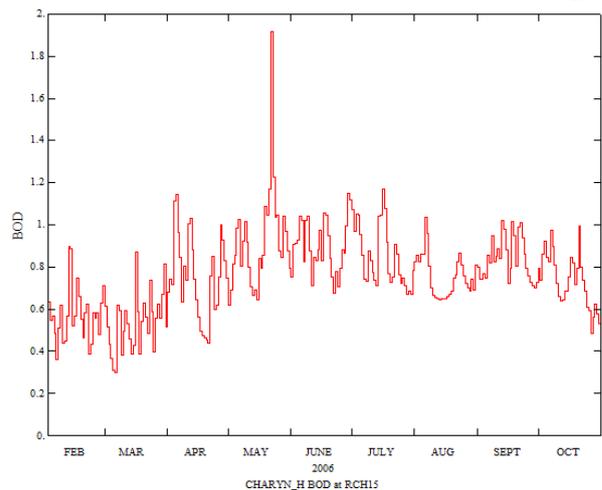


FIGURE 10 Changes in the computed concentrations of BOD

4.2 DISSOLVED OXYGEN

According to the observation data for the year 2006 obtained at the station located in Sarytogay tract, average monthly

dissolved oxygen concentration equals to 9.94 mg / L.

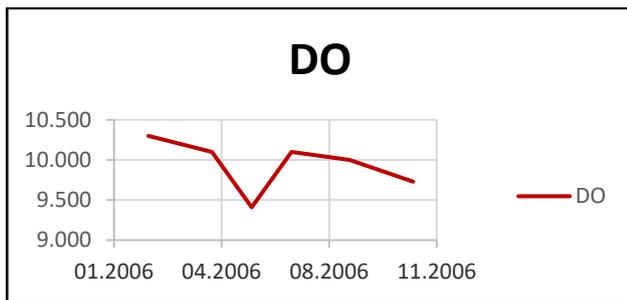


FIGURE 11 Changes in the concentrations of dissolved oxygen

Average annual concentration of dissolved oxygen at river reach R-15 obtained as a result of calculations made up 11.6 mg/L.

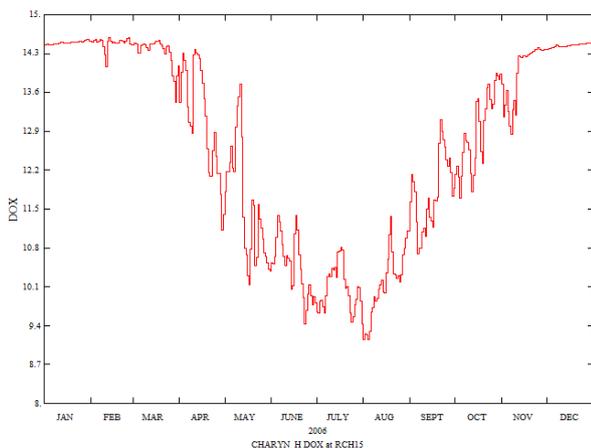


FIGURE 12 Changes in the computed concentrations of dissolved oxygen

4.3 NITRATE NITROGEN (NO3)

According to the observation data for the year 2006 obtained at the station located in Sarytogay tract, average monthly value of NO3 concentration = 0.58 mg/L.

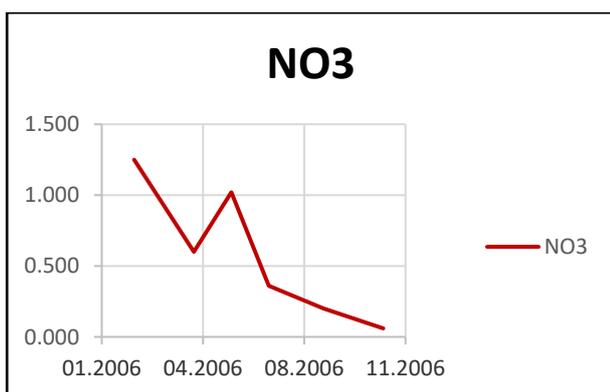


FIGURE 13 Changes in the concentrations of NO3

Average annual concentration of NO3 at river reach R-15 obtained as a result of calculations made up 0.51 mg/L.

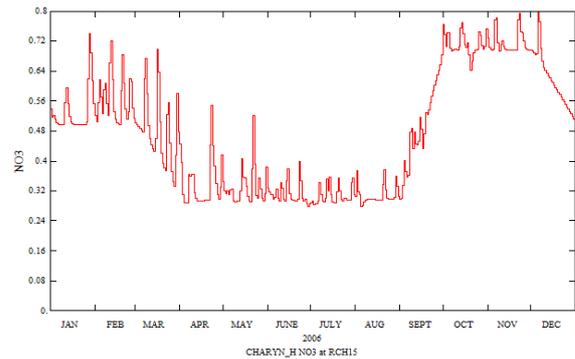


FIGURE 14 Changes in the computed concentrations of NO3

4.4 AMMONIUM NITROGEN (TAM)

According to the observation data for the year 2006 obtained at the station located in Sarytogay tract, average monthly concentration of TAM = 0.012 mg/L.

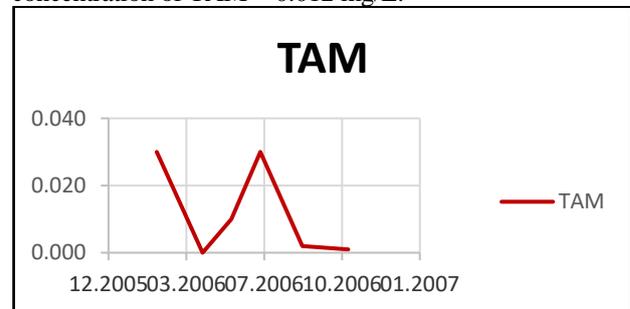


FIGURE 15 Changes in the concentrations of TAM

Average annual concentration of NO3 at river reach R-15 obtained as a result of calculations made up 0.01 mg/L.

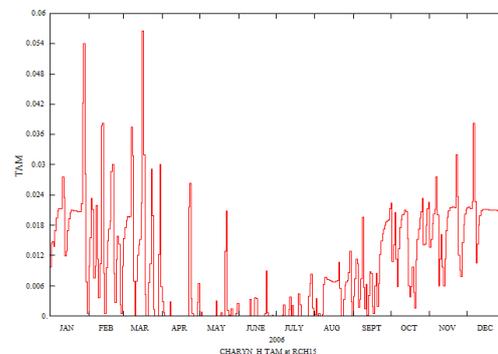


FIGURE 16 Changes in the computed concentrations of TAM

5 Conclusion

The main result of this work is approbation of BASINS software by the example of the modeling of transport pollution in small River Charyn in the south-eastern part of the Republic of Kazakhstan (Central Asia). The practice has shown that a simple installation of the product is not enough for the calculations. It was necessary to adapt the model to local conditions. However, some input parameters of the model had to be assembled from a variety of sources, transformed and adjusted.

Also, to verify the adequacy of the model, existing data were searched and an expedition with data

collection by field studies were carried out.

Comparisons on the following parameters:

- 1) Deviation of BOD value - 0.09 mg/L, error - 11%
- 2) Deviation of DO value – 1.66 mg/L, error - 16%
- 3) Deviation of NO₃ value - 0.07 mg/L, error - 12%
- 4) Deviation of TAM value - 0.002 mg/L, error - 16%

It can be noted that the model adequately reflects the situation on the transport of pollutants. Accuracy is the permissible value for the assessment of the environmental situation in the region.

In future will be planned to expand the usage of this model for large river basins, to consider the possibility of

calculation of point sources taking into account the daily load. After multiple calculations are also planning to create an automated system using distributed computing, similar to the [7].

Acknowledgements

We would like to express a gratitude to reviewers for valuable comments. The work supported by the grant No. 1746/GF4 of Ministry of Education and Science of The Republic of Kazakhstan.

References

- [1] «EPA's Better Assessment Science Integrating Point and Nonpoint Sources (BASINS)» *Exposure Assessment Models* <https://www.epa.gov/exposure-assessment-models/basins>
- [2] Andrew Battin, Russel Kinerson, Ph.D., Mohammed Lahlou, Ph.D. EPA's Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) *A Powerful Tool for Managing Watersheds 2014*
- [3] *Global Data Explorer* <http://gdex.cr.usgs.gov/gdex/>
- [4] *U.S. Releases Enhanced Shuttle Land Elevation Data* <http://www2.jpl.nasa.gov/srtm/>
- [5] «The European Commission's science and knowledge service» *Joint Research Centre* <http://forobs.jrc.ec.europa.eu/products/glc2000/glc2000.php>
- [6] *FAO GeoNetwork* <http://www.fao.org/geonetwork/srv/en/main.home>
- [7] Zaurbekov D L, Bostanbekov K A, Jamalov J K, Kim D K, Nurseitov D B, Tursunov I E, Zakarin E A 2013 Service-Oriented GIS System for Risk Mapping of Oil Spills Integrated with High Performance Cluster *The Second International Conference on Informatics Engineering & Information Science (ICIEIS2013)* 343-354

AUTHORS	
	<p>Jalal Jamalov, 02.03.1991, Kazakhstan</p> <p>Current position, grades: PhD student of Kazakh National Technical University. University studies: Master degree in computer engineering and software from International University of Information Technologies. Scientific interest: Computer modeling and ecological problems. Publications: 1. Water resources of Central Asia and their use. Materials International Scientific-Practical Conference devoted to the summing-up of the "Water for Life" decade declared by the United Nations. Volume 1. Almaty, Kazakhtan, September 22–24, 2016. 2. Numerical solution of the inverse problem Streeter-Phelps closed system for two incubation periods// Water resources of Central Asia and their use. International scientific-practical conference dedicated to summing up the results of the UN Declaration of the Decade "Water for Life". Book 1. Almaty, Kazakhstan, 22-24 September 2016 p. 56-63. 3. Example simulation of water protection measures in the Kazakh part of the basin. Or (scenario calculation) // Russian Foundation for Basic Research, Institute of Arid Zones, Southern Scientific Centre of the Russian Academy of Sciences, Southern Federal University. Collection of articles "Ecology, Economics, Computer Science. System analysis and modeling of economic and ecological systems. Mathematical methods and models in environmental studies." Volume 1. Rostov-on-Don, 2016. p. 33-43. Experience: Experience in software development, web applications and GIS technologies.</p>
	<p>Daniyar Nurseitov, 22.03.1980, Kazakhstan</p> <p>Current position, grades: Head of National Open Research Laboratory Information and Space Technologies at Kazakh National Research Technical University. University studies: Candidate of physical and mathematical sciences from Novosibirsk State University. Scientific interest: Inverse problems, computer modeling and ecological problems. Publications: 1. Analysis of ill-posedness and numerical methods of solving a nonlinear inverse problem in pharmacokinetics for the two-compartmental model with extravascular drug administration // Journal of Inverse and Ill-Posed Problems. Volume 20, Issue 1, P. 39–64, March 2012. 2. An optimization method in the Dirichlet problem for the wave equation // Journal of inverse and ill-posed problems. Volume 20, Issue 1, Number 2, P. 193–211. 3. Inverse problems for the ground penetrating radar // Journal of Inverse and Ill-Posed Problems. Volume 21 Number 6, P. 885–892, 2013 Experience: Leadership experience five successful scientific projects.</p>
	<p>Kairat Bostanbekov, 05.07.1988, Kazakhstan</p> <p>Current position, grades: PhD student of International University of Information Technologies. University studies: Master degree in computer engineering and software from Kazakh National Technical University. Scientific interest: Computer modeling and ecological problems. Publications: 1. Integrated workflow-based system for risk mapping of oil spills with using high performance cluster // International Journal of New Computer Architectures and their Applications (IJNCAA): Vol. 3, Issue 4. P.115–131, 2013. 2. Service-Oriented GIS System for Risk Mapping of Oil Spills Integrated with High Performance Cluster // The Second International Conference on Informatics Engineering & Information Science (ICIEIS2013):Kuala Lumpur, Malaysia, Nov. 12-14, P.343-354, 2013. 3. Geoinformacionnaia model bioraznobrazia i uyazvimosti bioty Severnogo Kaspia [Geoinformation models biodiversity and vulnerability of biota of the Northern Caspian] // Geoinformatika P.55-63, 2016. Experience: Experience in software development, web applications and GIS technologies.</p>