Jan Jędrasik, Marek Kowalewski, Bogdan Ołdakowski, Jacek Namieśnik

Calibration, verification and validation of the Baltic Ecohydrodynamic Model

Literature
Acknowledgements

Abstract

The present article is a continuation of the paper I, entitled "The ecohydrodynamic model for the Baltic Sea". The introduced changes in the algorithm of the ecosystem part – the ProDeMo model describing environmental processes - have entailed necessity of new procedure steps: calibration, verification and validation. New assumptions were taken in present approach of calibration. The first was horizontal distinction of aquatories as three areas such as coastal, open gulf and open sea. Next - dividing waters into three layers: surface, intermediate and bottom in vertical. The other assumption concerned the time stages from three to nine years using two different resolutions. The interactive database was created to investigate and compare effectiveness of the new simulation to basic one. Each effectiveness referred to simulations and observations of three state variables groups: nutrients, total forms of nitrogen and phosphorus as well as for dissolved oxygen. The improvement of effectiveness coefficient in sequential simulations referring to basic run indicated the progress of calibration. As a result of new calibration 96 coefficients were obtained, which completed the equations describing biogeochemical processes. They are attached in the paper I.

The ecohydrodynamic model was run for different length of cycles: one (1995), three (1994 – 1996) and nine (1994 – 2002) years. The annual cycle of 2002 was repeated for 15 years. All the computations verified that model as a convergent, stable and consistent one with the mass conservation law. For the last procedure step – validation - simulations coming from new calibrated model have been run for real meteorological and hydrological conditions for period 1994 – 2002 and were compared to observations. Statistical measures used for estimation of the simulations indicate the high quality of the model.

1. Introduction

Mathematical modelling is a method of research which enables quantitative and qualitative analysis of processes taking place in natural environment. The paper entitled "The ecohydrodynamic model for the Baltic Sea" (as the paper I) presents such an approach. This approach also allows to estimate influence man's activity or the prognosis of future changes in ecosystem (Kannen et. al., 2004).

The presented ecohydrodynamic model has been developed (Ołdakowski, Kowalewski, Jędrasik, 1994; Oładkowski and Renk, 1997; Kowalewski, Jędrasik, Oładkowski, 2003). Last version (Kowalewski, et al., 2004) was applied for investigation of biogeochemical processes at the water environment of the Gulf of Gdansk for period 1994 – 2002.

Former calibration of the ProDeMo model (Kowalewski, Jędrasik, Ołdakowski, 2003) provided the set of complementary coefficients to equations describing biogeochemical processes at the southern Baltic with particular stress on the Gulf of Gdansk environment. The method was based on a comparative analyses of sequential model results as well as analyses of sensitiveness of state variables. Another advantage of that calibration was using statistic measures for evaluation of effectiveness modelled simulation in comparison with the measurements. However, that approach was found as time-consuming.

Recent development of biogeochemical processes description caused essential changes of the model algorithm. There were as following: extension number of phytoplankton groups to five, including nitrogen fixation process, splitting of nutrients deposition on two phases: their accumulation into active layer and burring in the deeper passive one.

Due to the introduced changes in the algorithm, the necessity to pass a new procedure steps: calibration, verification and validation of the model ProDeMo was found. Drawing a conclusion from the former calibration of the model and using the statistics measures, it has been elaborated new favourable calibration. That approach resulted in new set of coefficients to be completed the equations biogeochemical processes. The outcome obtained from comparison between model simulations and measured data as high coefficients of effectiveness gave evidence that the processes have been described accurately as well as the model has got better quality. However, the model was subjected to validation using measurements taken from coast waters in front of the mouth of the Vistula river. The obtained compliance with the observations confirmed that the model was a tool for reliable prediction of ecosystem behaviour in the southern Baltic.

2. Methodology of research

Practical building a model with new algorithm demands sequential procedure steps (Fig. 1) which allow to find the best accordance between computed and observed data by fitting of relevant coefficients in calibration. It is rather an iterative operation. To test an internal logic of the model, if the model react as expected, what is its behaviour - a verification is needed. Next, investigation how well the model simulations fit the in situ data describes a validation step.

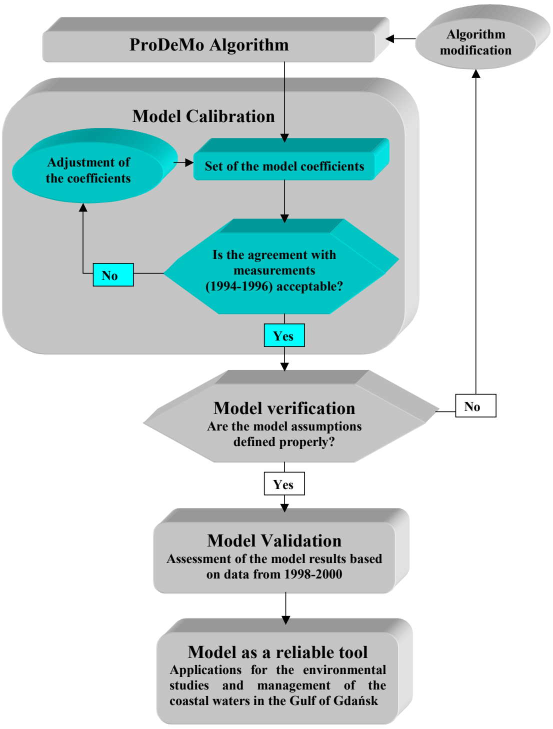


Fig. 1 Scheme of the model procedure steps

2.1 Assumptions for the calibration of the model

Based on the experience, new assumptions for the calibration have been made. Assessment of the model quality ought to be based on statistical measures, particularly on the effectiveness coefficient evaluated between observed and modelled values of state variables (Appendix I). The basic compared quantities were the modelled and observed values. The differences between them were defined as **a model error,** which has risen to the square became the **mean quadratic error.** To investigate what degree the modelled values were underestimated or overestimated in relation to the observed ones, difference of averages called **absolute bias** were calculated**.** The correlation coefficient as a product of calculated and observed standard quantities as well as the standard deviation of the differences between observed and modelled values were evaluated. Relationship between correlation coefficient and biases contained in the mean quadratic error, allowed to deduce the **coefficient of effectiveness,** which was assumed as a basic one for optimisation of model calibration. Another important statistical measure was the **special coefficient of correlation** which was a result of the relation between the correlation coefficient and a **total quadratic error.** This coefficient was assumed for estimation of the model's quality. These formulas were taken as an algorithm in calibrating database for judgment of the simulations during calibration's processes.

The assumptions included comparison of sequential simulations to basic one. All of them referred to the same set of measured data. Simulations in the next runs of the model ought to be affected by changes of calibrating coefficients.

Distinction of marine aquatories was another assumption during investigation of the simulations. There were three areas: coastal, open boundary of the Gulf of Gdansk and the southern Baltic that have been taken into account. Stations: ZN2, NP, K and P101 which were located in front of the mouth of the Vistula river belonged to the coastal area. The open boundary waters contained P1 and P110 stations, the others: P5, P39, P63 and P140 were located at the third area (Fig. 2).

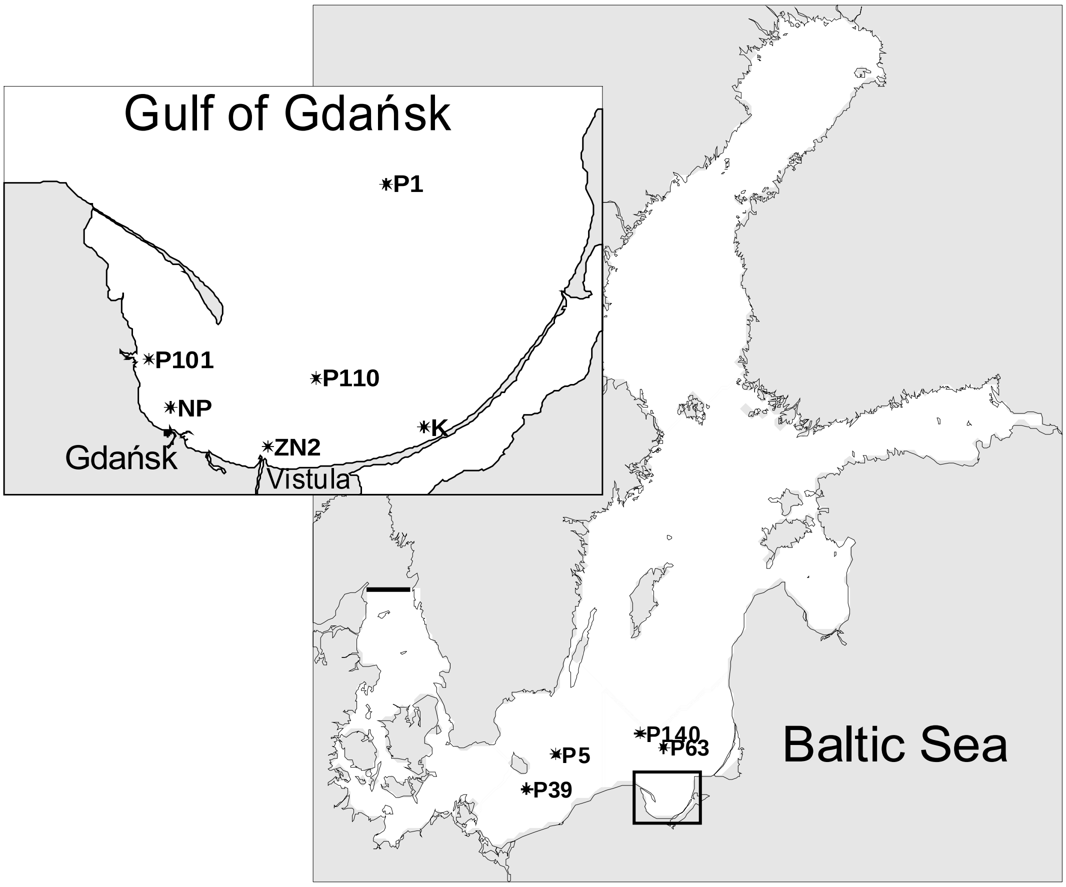


Fig. 2 Modelled areas with marked observation stations

Analyses of effectiveness were also made in vertical direction dividing waters into three layers: surface (0 – 30 m), intermediate (30 – 70 m) and bottom ones (beneath 70 m). Calibration's processes need to be based on time stages - three and nine years ones – using different horizontal resolutions for the Baltic and the Gulf of Gdansk. The first period (1994 - 1996) and the second one (1994 – 2002) referred to the model of the whole Baltic. The third period (1994 – 1996) concerned both areas with their resolutions.

2.2 Calibration's description

The interactive database in the computer program access was formulated due to mentioned assumptions. It referred to both areas, the Baltic and the Gulf of Gdansk, for period between 1994 – 2002. The database contained of measured data and modelled simulations. The effectiveness of the new simulation was analysed and evaluated for three groups of state variables: nutrients (NO_3, NH_4, PO_4, SiO_4), total forms of nitrogen and phosphorus (N_{TOT}, P_{TOT}) as well as for oxygen dissolved (O_2) (Tab. 1).

Table 1. Coefficients of effectiveness for nutrients (E_B), total forms of N,P (E_{TOT}) and oxygen (E_{O2}) for basic and final model

effectiveness	**Basic**	**Final**	**Change**
$\mathbf{E_B}$	0.37	0.42	0.05
E_{TOT}	0.20	0.31	0.10
E_{O2}	0.74	0.76	0.02

The effectiveness was also estimated for each state variables separately as well as for temperature T and salinity S (Tab. 2).

Table 2. Coefficients of effectiveness of state variables for basic and final model

Parameter	**Basic**	**Final**	**Change**
NH4	0.29	0.32	0.03
NO3	0.53	0.52	-0.01
NTOT	0.00	0.05	0.05
O2	0.74	0.76	0.02
PO4	0.46	0.58	0.12
PTOT	0.41	0.57	0.16
S	0.82	0.79	-0.03
SiO4	0.21	0.28	0.07
T	0.91	0.92	0.00

The following step was the research of effectiveness in layers. It was compared between the basic model and the present one. It was observed clearly which variables were simulated better or worse (Fig. 3). The effectiveness coefficients for compared the basic and present models were illustrated on the left panel and differences between them on the right one.

Some additional analyses have been undertaken to change the effectiveness coefficient for new simulation. The bias were considered for state variables to evaluate whether they were underestimated or overestimated in relation to the observed ones. The simulation was compared to observations at the chosen depth of a particular station. The next steps were following: the observation of the simulated sequences of five groups of phytoplankton blooms, primary production, phyto- and zooplankton sedimentation as well as detritus. Then both the rate of deposition and the structure of sediments, the results of fixation and nitrification processes have been analysed. Due to this analyses it was indicated which coefficients ought to be modified before next simulation. Each simulation provided arguments whether the progress appeared comparing to the basic model. The necessity of introducing the new basic model has occurred to obtain better results of calibration.

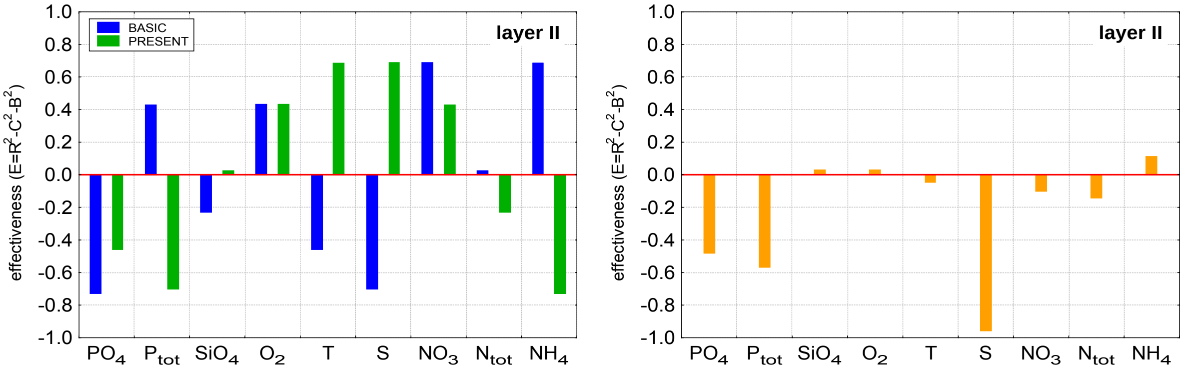


Fig. 3 Coefficients of effectiveness for state variables at layer II in 1994 – 2002,

2.3 Stages of the calibration process

The first stage of calibration referred to the whole Baltic with simulations for 1994 – 1996. Obtained coefficients presented changes in phytoplankton, nutrients and oxygen. Decreasing of maximum temperature of phytoplankton growth (T_{max}) for dinoflagellates improved the simulation. For blue green algae the optimum and minimum temperatures (T_{opt} T_{min}) as well as sedimentation rate (V_s) were decreased. However, optimum light intensity (I_s) was increased. The maximum temperature has been decreased and optimum light intensity was increased for autumn diatoms. Their phytoplankton mortality rate (L) slightly increased. For green algae optimum light intensity was increased. The nitrification coefficient (K_{nN}) has been increased for nitrogen. For phosphorus the phosphorus mineralization rate in water (K_{mP}) and sediments (K_{mPS}) were increased but the silicon mineralization rate (K_{mSi}) was decreased. The parameter for oxygen flux in case of over saturation conditions (B_{DO}) was increased what implied for decreasing of the reaeration coefficient (B_{DOW}).

In the second stage of calibration the simulations were extended from three to nine years (1994 – 2002) and still referred to the whole Baltic area. Sequential changes of coefficients referred to the phytoplankton contained increasing of the sedimentation rate for both the spring diatoms and the green algae as well as the optimum light intensity for autumn diatoms. Considering a slow growth of the blue green algae, the parameter of food availability (P_{aval}) has been increased. The phosphorus and silicon mineralization rate were lowered in the water but has been risen in the sediments.

The third stage comprised both the Baltic and the Gulf of Gdansk for period of three years (1994-1996). Despite the differences in resolutions of numerical grids the same coefficients were used for both areas. Except for zooplankton coefficients the changes have been made in each group of remaining ones. The sedimentation rate for the spring diatoms has been decreased. For blue green algae the optimum and minimum temperatures as well as the parameter of food availability were decreased. The optimum light intensity rose for dinoflagellates and blue green algae. Oxygen parameter for carbon mineralization rate was decreased while the nitrification coefficient increased. Quite a number of changes have been registered for phosphorus e.g. the phosphorus mineralization rate in water and sediments also fraction of phosphorus adsorbed on inorganic particles have been lowered. The parameter for oxygen flux in case of over saturation conditions was decreased, too. Oxygen to carbon ratio during photosynthesis (a_{OC}) grew up.

On account of long computing time the forth stage of calibration (1994 – 2002) including both areas, it was only initiated. The phytoplankton calibration coefficients were identical as at the previous stage. The only differences referred to values for sedimentation rate of dinoflagellates as well as silicon mineralization rate for the autumn diatoms. Moreover, the detritus sedimentation rate (Vs_{DE}) was decreased but the nitrogen mineralization rate in sediments (K_{mNS}) rose.

There were adequate numbers of simulations for each particular calibration stage: 42 (first stage - I), 11 (second stage - II), 4 (third stage – III) 1 (forth stage - IV). Increasing or decreasing of three effectiveness coefficients $\mathbf{E_B}$ **(nutrients),** $\mathbf{E_C}$ **(**total forms of N,P**)**, $\mathbf{E_{O2}}$ (oxygen) was a result of its comparison at the beginning and the end of each calibration stage (Tab. 3). A significant improvement for $\mathbf{E_C}$, $\mathbf{E_{O2}}$ and a slight one occurred during the first stage. The second widen calibration stage effectiveness for oxygen was constant, little improved for total forms but slightly decreased for nutrients. The calibration performed for both areas with different resolutions of numerical grids, caused lower effectiveness. However, its little improvement appeared for oxygen and total forms. Due to this calibration obtained a new set of complementary coefficients (Appendix II in paper I) to equations describing biogeochemical processes at the southern Baltic with particular stress on the Gulf of Gdansk environment.

Table 3. Coefficients of effectiveness for stages (I – IV) of the model's calibration

Effectiveness	**I**		**II**		**III**		**IV**	
	Basic	**Final**	**Basic**	**Final**	**Basic**	**Final**	**Basic**	**Final**
$\mathbf{E_B}$	0.37	0.38	0.38	0.36	0.28	0.27	0.28	0.25
$\mathbf{E_C}$	0.20	0.45	0.45	0.48	0.18	0.15	0.18	0.35
$\mathbf{E_{O2}}$	0.74	0.79	0.79	0.79	0.68	0.74	0.68	0.76

Most often published scientific literature presents similar investigations limited to one or few stations. This calibration was based on all measurements at the available monitoring stations during period 1994 – 2002. Due to such a number of measurements the results of calibrations coefficients are statistically essential. They contribute to the increase of the model's simulations quality. This kind of calibration is an important procedural step in ecological modelling and renders the model as a reliable scientific research tool.

3. Verification of the model

Solution of the equations completed by calibration coefficients has been given a reasonably realistic values. The model was run during annual cycle, next repeated for three years and finally for nine years periods. At these stages concentration of nitrogen and phosphorus compounds, phytoplankton biomass followed by due to fluctuating atmospheric forcing and variable of nutrient loads. The simulated parameters were the same scale as observed both in vertical as well as horizontal distributions (Fig. 4a - b; 5a - c). The model repeated seasonal changes of primary production. It has appeared as long term stable even for 15 years forecast (2000 – 2015) and complied with the mass conservation law. These features certify its internal logic and correctness.

4. Comparison of the simulations versus measurements

Here are the results after calibration process to define the relationship between observations and simulations of the selected variables: concentrations of nitrate nitrogen [N-NO3], ammonium nitrogen [N-NH4], total nitrogen [N-Tot], phosphate phosphorus [P-PO4], total phosphorus [P-Tot], silicate silicon [Si-SiO4], dissolved oxygen [O2], temperature [T] and salinity [S] of water in the southern part of the Baltic Sea mainly at the Gulf of Gdansk. The modelled values were compared with those measured on standard levels of depth in selected observation stations.

4.1 Vertical distribution of variables in seasonal formulation

Vertical distribution of selected calculated and measured parameters were analysed on the example of station P1 for two seasons winter and summer in 2000 (Fig. 4a, b). The modelled values of the vertical distribution both nitrate nitrogen and ammonium nitrogen were overstated.

Phosphate and total phosphorus simulations down to 70 m (halocline) were described well and those below slightly underrated (Fig. 4a). Silicate silicon values modelling for the winter period showed convergence with the observations down to 80 meters but they were underrated in the bottom layers. The best described distribution of dissolved oxygen was to the halocline depth – 70 m. Beneath measured values indicated for its budget deficit in opposite to model (Fig. 4a). The weak permanent thermocline was showed more clearly by measured winter profiles of the water temperature and salinity than model simulation (Fig. 4a).

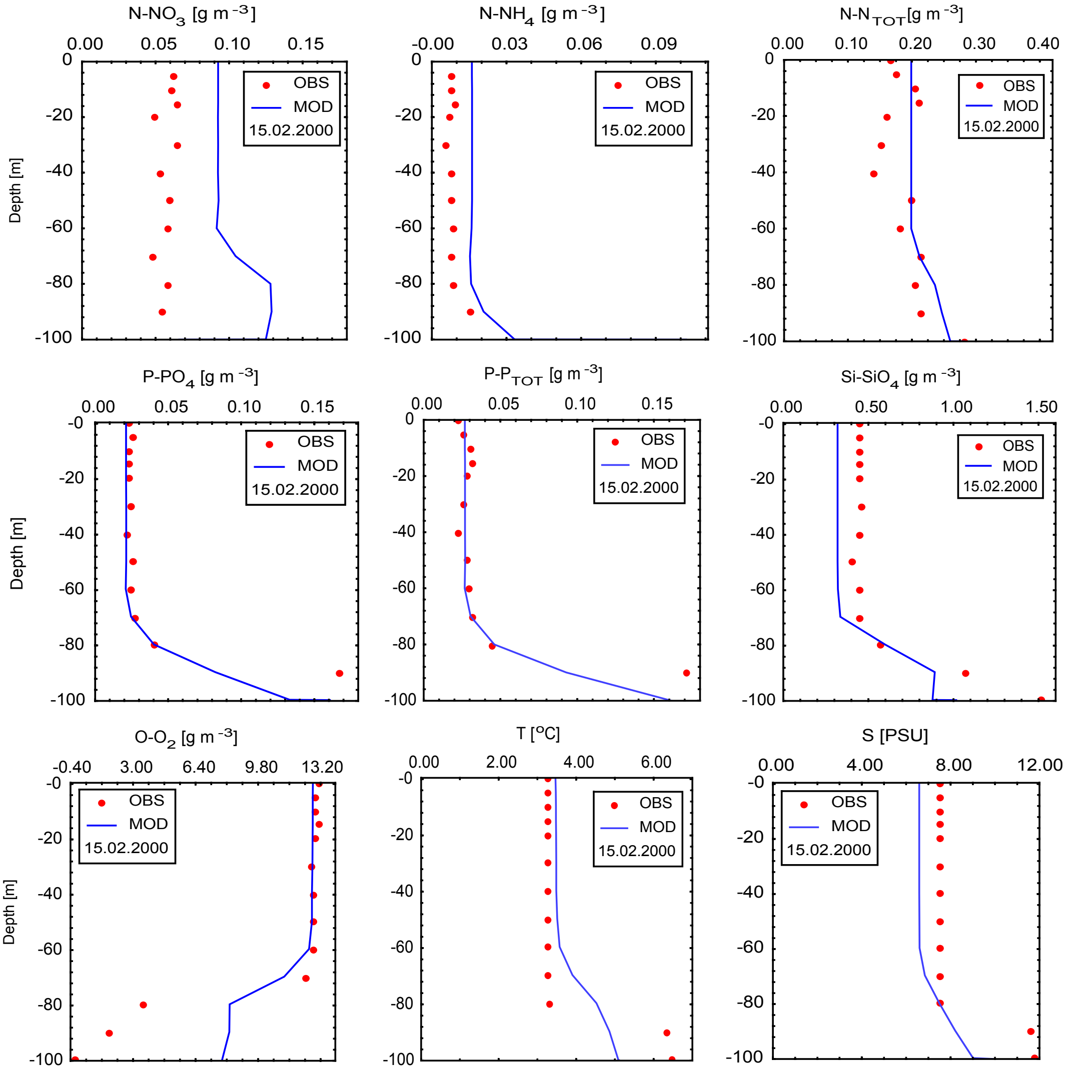


Fig. 4a. Winter variability in 2000 of vertical distributions of the observed (OBS) and modelled (MOD) chemical and physical parameters at the Gdansk Deep, station P1

In surface layer for summer the modeled nitrate concentrations were in accordance with the measurements. The summer period observations indicated, however, that model simulations indicated much shallower nitrate depletion than the observations would show. Their full depletion down to 60 meters while the model simulations only down to 20 meters (Fig. 4b). The calculated distribution of ammonium nitrogen concentration in particular season of 2000 was in conformity with the measured values. The phosphate phosphorus concentrations reached the values approaching those measured during the summer in the layer to 70 meters. Comparing vertical distribution of total forms of nitrogen to phosphorus has been noticed much favourable in the water column for phosphorous. It was similar with silicate silicon and dissolved oxygen. In the bottom layer below halocline the calculated values of concentrations of silicates and phosphates were underrated, and concentration of oxygen was overstated in relation to the observed ones. Vertical distributions of salinity demonstrated better consistence with observations than water temperature (Fig. 4b).

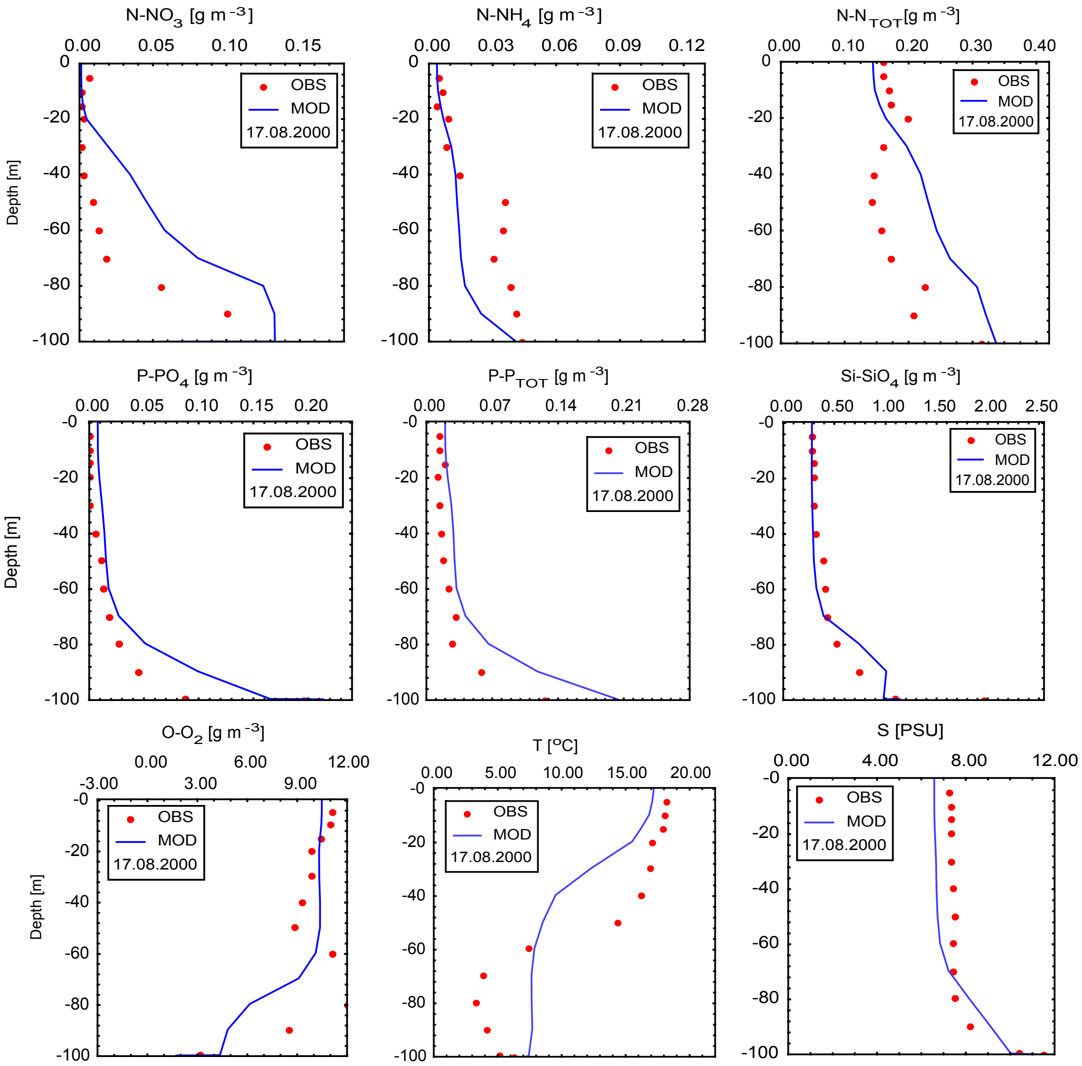
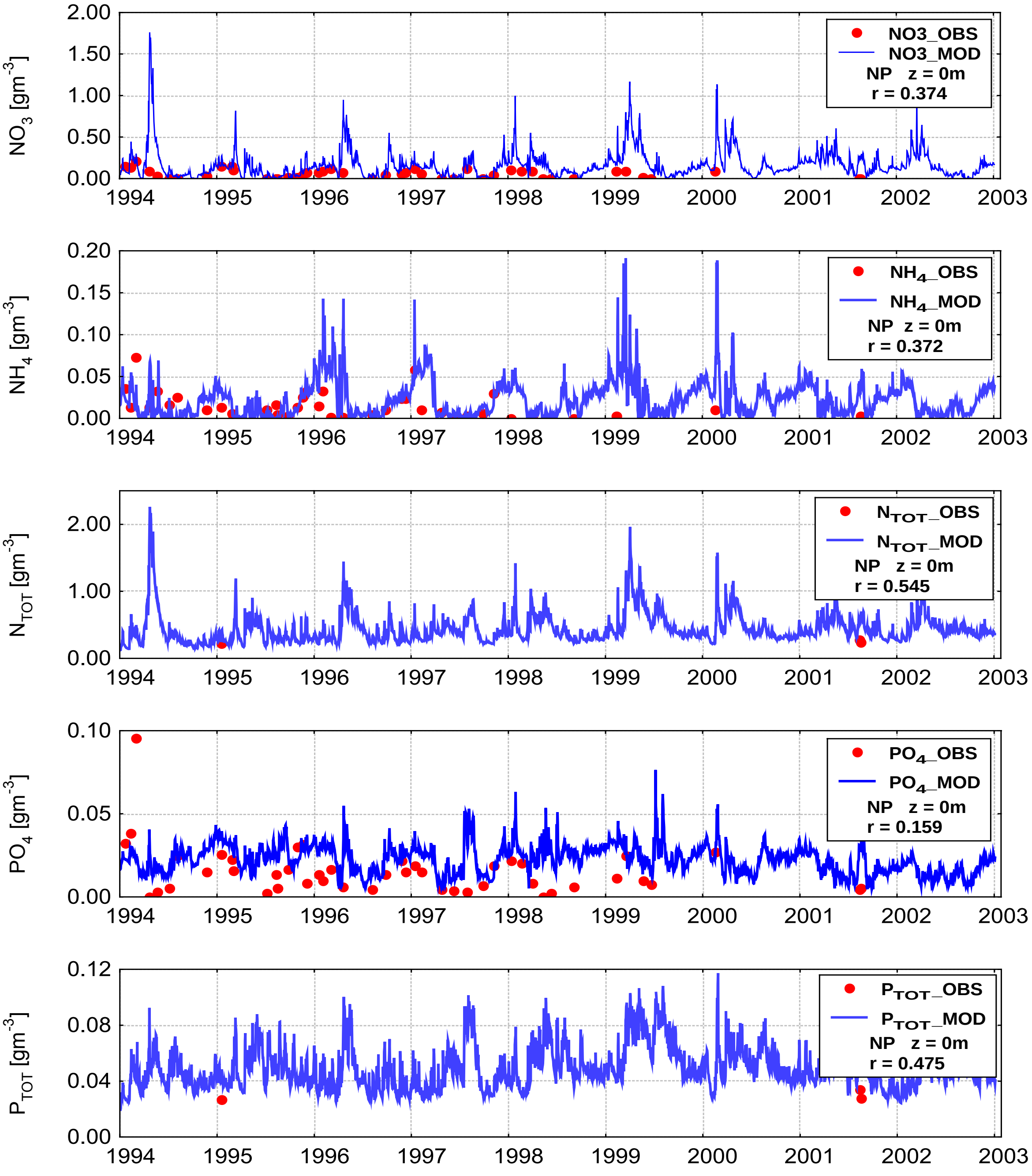


Fig. 4b. Summer variability in 2000 of vertical distributions of the observed (OBS) and modelled (MOD) chemical and physical parameters at the Gdańsk Deep, station P1

Apart from vertical variability, surface runs of parameters were analyzed. Modelled and measured surface values of state variables have been compared at mentioned stations as well as the bottom one (z = 100 m) at the station P1 (the results fro station P1 are showed in the paper I). The best correlations between simulations and observations at the surface station P1 for temperature (r =0.97), oxygen (r = 0.85) phosphates (r = 0.76) and nitrates (r = 0.66) were obtained. In contrary ammonium and silicon were modeled inadequately (r = -0.06 and r = 0.15). Highest correlation of coefficients were received as follows: nitrates (r = 0.68), ammonium (0.53), phosphates (r = 0.0.76) silicon (r = 0.21), oxygen (r = 0.85) temperature (r = 0.98) and salinity (r = 0.63) adequately at stations K, ZN2, P1, ZN2, P1 and NP (Fig. 5a – c). The parameter which has been best described by the model up till now is the sea water temperature. The values of the ammonium and silicon differentiated between the stations P1 and ZN2 (at the 50 km distance) by one order (Fig. 5a and c). It was observed similar relationship for total nitrogen and phosphates. The salinity was the most stable parameter at the station P1 and very highly changeable at the ZN2 (because of the Vistula river influence).

In the Gdansk Deep at station P1 the analysis of parameter fluctuation was made at the bottom on 100 meters. Correlation for each variables fell regularly e.g., for temperature from r = 0.98 at the surface to r = 0.19 at the bottom. Observations showed the deficit of dissolved oxygen while simulations do not describe this phenomenon. Despite of that, the oxygen was the best modeled parameter at the bottom (r = 0.35). The next important feature was underestimating of values especially in the deeper layers, which is confirmed by vertical distributions (Fig. 4a, b).

Comparison of variables at stations NP, ZN2 and K during 1994–2002, demonstrated recurrence of annual cycles without any definite trend (Fig. 5a-c). This testifies good functioning of the model. Each of the researched stations showed a regularity of the summer depletion of mineral forms of nitrogen and phosphorus. In this period silicate silicon exhibited a lowered level of concentration. Variability of oxygen, which was in counter phase to water temperature showed characteristic cycle. The rise of temperature corresponded to lowered values of dissolved oxygen concentrations. It should be emphasized that the values modeled for the period 1994–2002 departed from the measurements reasonably.



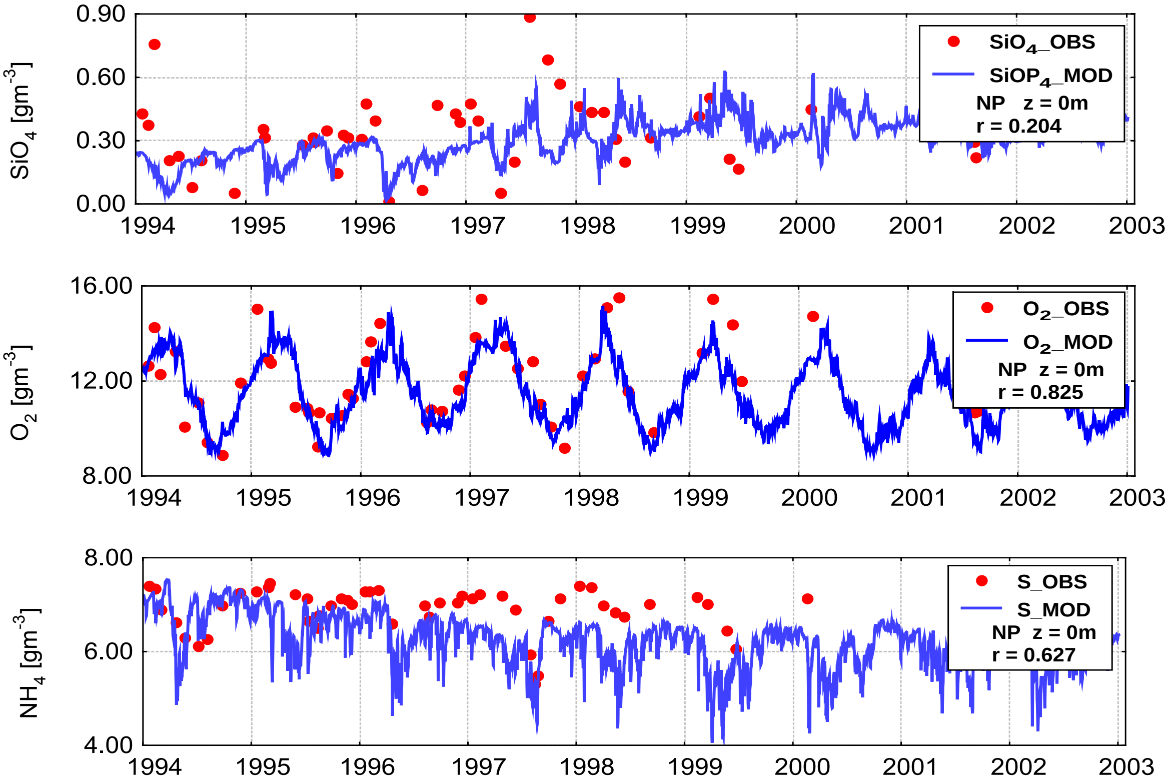
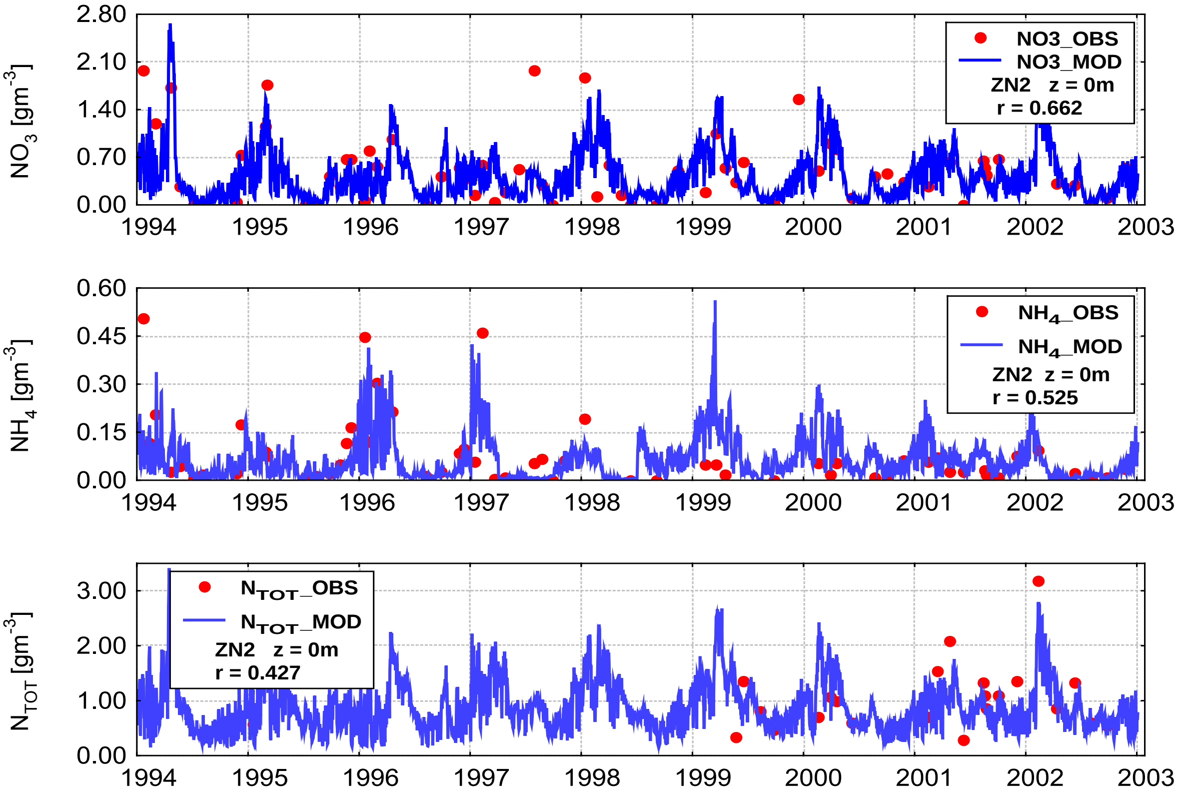
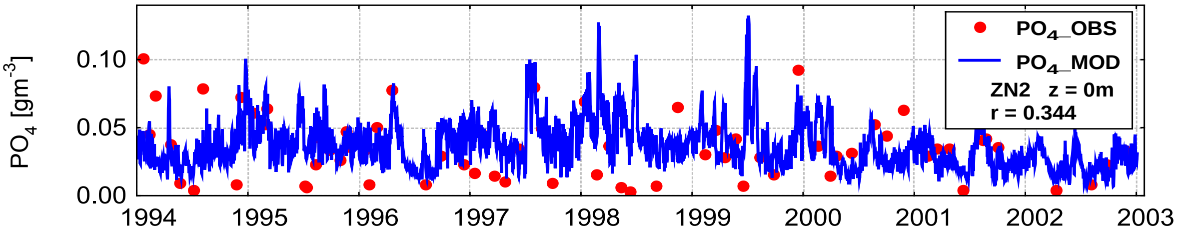
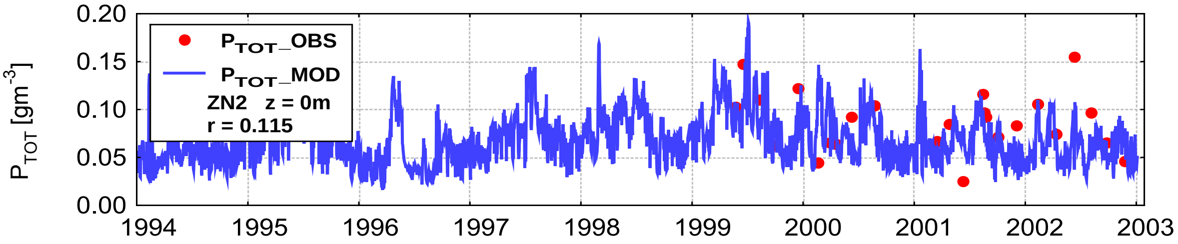
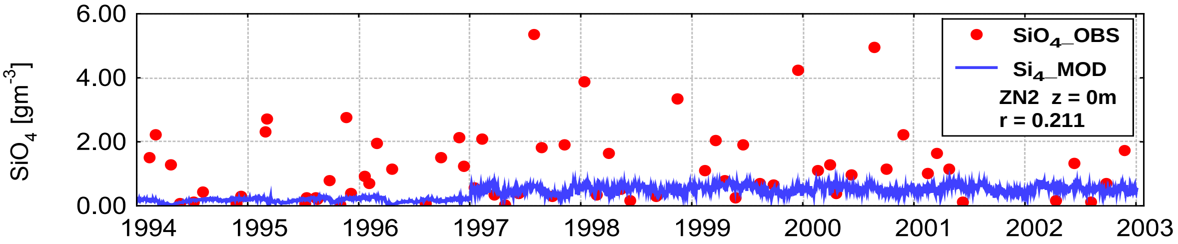


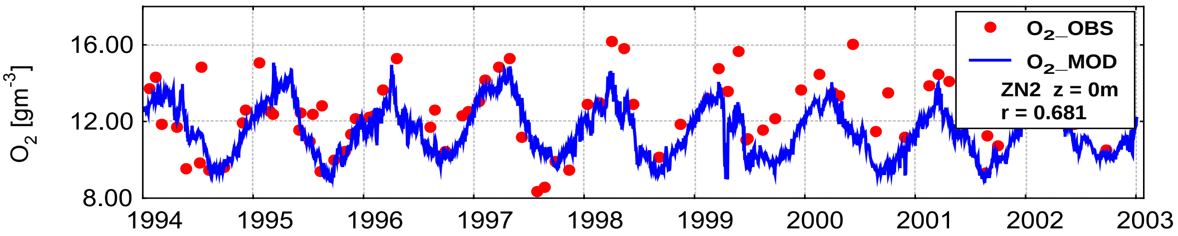
Fig. 5a The run of surface variability of the observed (OBS) and modelled (MOD) chemical and physical parameters: nitrates NO_3, ammonia NH_4, total nitrogen (N_{TOT}), phosphates PO_4, total phosphorus (P_{TOT}), silicates SiO_4, dissolved oxygen O_2, water temperature T_w (station **NP**, depth **0** m) in 1994 – 2002

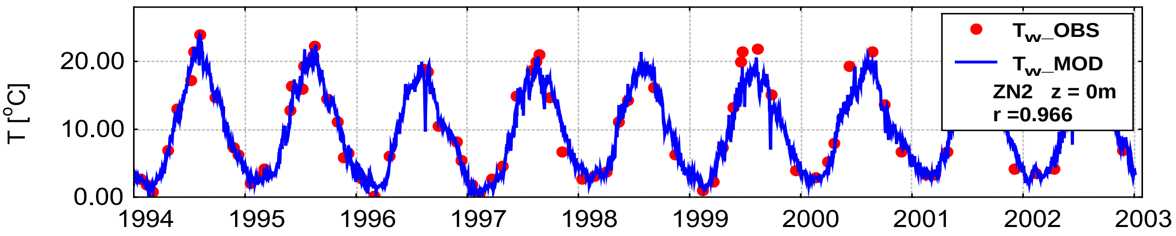



PO$_4$ [gm^{-3}]
0.00
0.05
0.10
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
PO$_4$_OBS
PO$_4$_MOD
ZN2 z = 0m
r = 0.344


P$_{TOT}$ [gm^{-3}]
0.00
0.05
0.10
0.15
0.20
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
P$_{TOT}$_OBS
P$_{TOT}$_MOD
ZN2 z = 0m
r = 0.115


SiO$_4$ [gm^{-3}]
0.00
2.00
4.00
6.00
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
SiO$_4$_OBS
Si$_4$_MOD
ZN2 z = 0m
r = 0.211


O$_2$ [gm^{-3}]
8.00
12.00
16.00
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
O$_2$_OBS
O$_2$_MOD
ZN2 z = 0m
r = 0.681


T [^{o}C]
0.00
10.00
20.00
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
T$_w$_OBS
T$_w$_MOD
ZN2 z = 0m
r =0.966

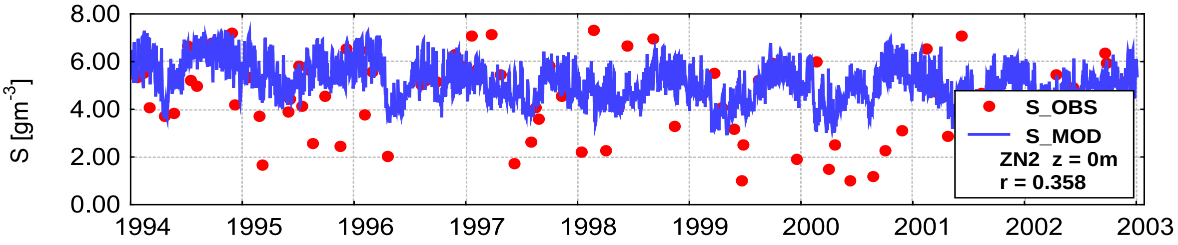
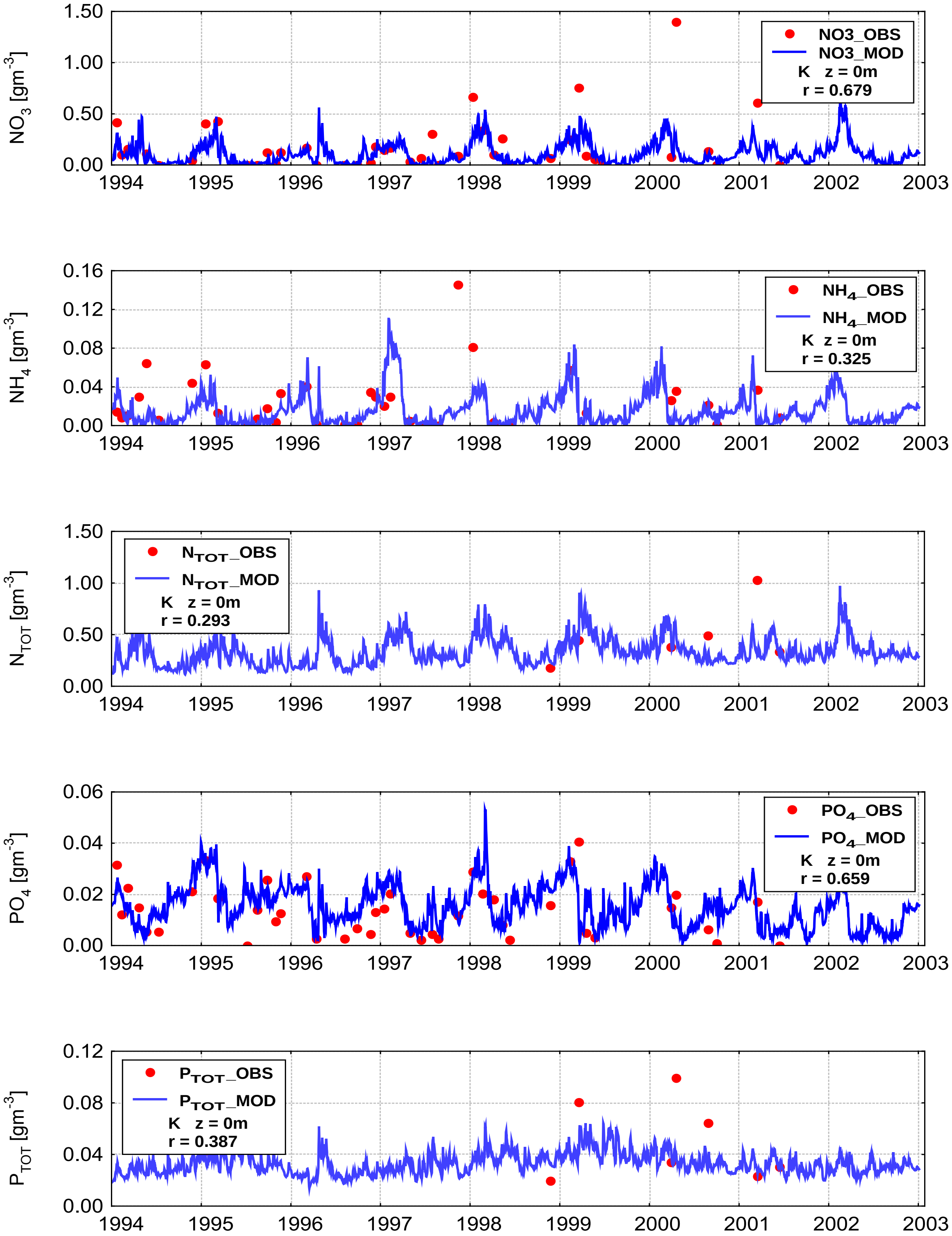

S [gm^{-3}]
0.00
2.00
4.00
6.00
8.00
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
S_OBS
S_MOD
ZN2 z = 0m
r = 0.358

Fig. 5b The run of surface variability of the observed (OBS) and modelled (MOD) chemical and physical parameters: nitrates NO_3, ammonia NH_4, total nitrogen (N_{TOT}), phosphates PO_4, total phosphorus (P_{TOT}), silicates SiO_4, dissolved oxygen O_2, water temperature T_w (station **ZN2**, depth **0** m) in 1994–2002



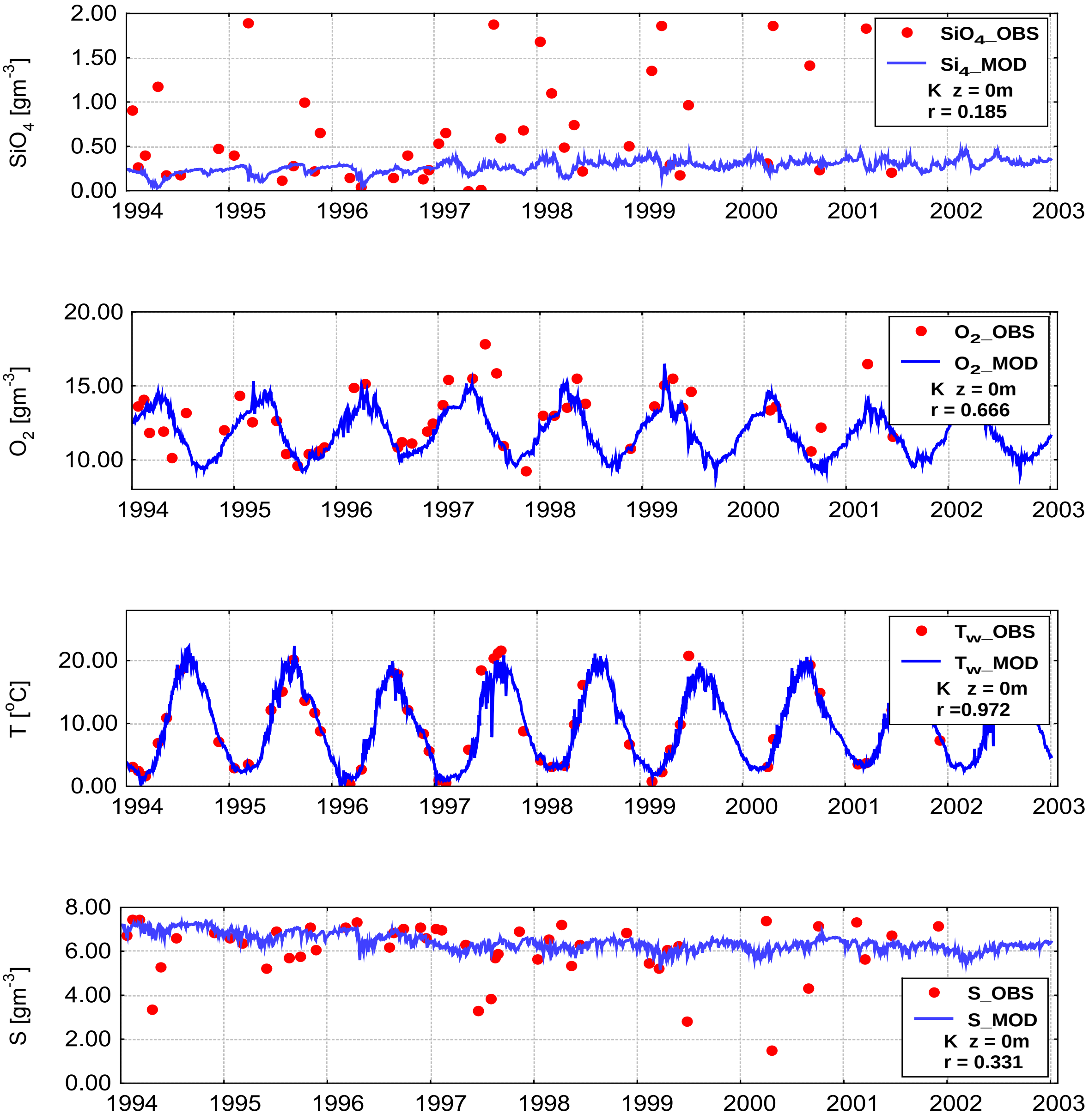


Fig. 5c The run of surface variability of the observed (OBS) and modelled (MOD) chemical and physical parameters: nitrates NO_3, ammonia NH_4, total nitrogen (N_{TOT}), phosphates PO_4, total phosphorus (P_{TOT}), silicates SiO_4, dissolved oxygen O_2, water temperature T_w (station **ZN2**, depth **0** m) in 1994–2002

To evaluate the simulation quality of the state variables the correlation and effectiveness coefficients as well as the bias were applied. The simulations from all stations were subjected to these measures. The parameter which has been best described in the whole profiles was the water temperature (Tab. 4), similarly as at the surface. The oxygen was another parameter that obtained high correlation coefficient (0.72 – 0.91). Modeled and measured values of phosphates and ammonium have agreed quite well at the station P1 and nitrates at the station K. Computed values of temperature, oxygen and salinity were decreased during modeling but remaining variables increased, extremely – nitrates. The physical parameters were simulated the most effectively. Among nutrients phosphorus compounds were modeled twice better than nitrogen ones. Much effort is needed to be done for modeling of ammonia. The correlation coefficients were referred to vertical and horizontal structure of the aquatoriums. The best correlation coefficients were received for variables at the intermediate layer but at the surface one appeared good effectiveness (Tab. 5). Nitrogen compounds were the most increased parameters at all layers. Taking into account simulations localized at divide areas, the excellent correlation coefficients have been got at the second area (area II) – waters at the open boundary the Gulf of Gdansk. The highest effectiveness of simulations was obtained at this area too (Tab. 6). The simulations were mostly increased at the near shore area (area I) in front of the mouth of the Vistula river.

Table 4 Coefficients of correlation, bias and effectiveness for state variables at the stations of the Gulf of Gdansk in 1994 – 2002

Station	**NH_4**	**NO_3**	**N_{TOT}**	**O_2**	**PO_4**	**P_{TOT}**	**S**	**SiO_4**	**T**
correlation coefficient									
P1	0.68	0.51	0.62	0.91	0.84	0.83	0.92	0.87	0.94
ZN2	0.45	0.59	0.44	0.72	0.36	0.44	0.46	0.22	0.95
K	0.39	0.64	0.66	0.90	0.81	0.81	0.87	0.59	0.96
All ZG	0.40	0.63	0.64	0.88	0.74	0.70	0.90	0.59	0.96
bias									
P1	1.28	0.65	1.04	1.03	1.02	1.01	0.91	0.79	0.99
ZN2	1.54	1.73	1.47	0.98	1.52	1.25	0.9	0.51	0.98
K	1.01	1.65	1.17	1.01	1.16	1.07	0.91	0.74	0.99
All ZG	1.05	1.46	1.06	0.99	1.09	0.97	0.92	0.73	0.98
effectiveness									
P1	0.13	0.34	0.37	0.75	0.71	0.69	0.60	0.64	0.89
ZN2	-0.14	0.02	-0.28	0.49	-0.34	-0.01	0.06	-0.10	0.96
K	0.04	0.08	0.15	0.75	0.64	0.65	0.58	0.26	0.92
All ZG	0.04	0.16	0.20	0.76	0.55	0.49	0.70	0.24	0.91

Table 5 Coefficients of correlation, biases and effectiveness for state variables at the layers of the Gulf of Gdansk and southern Baltic in 1994 – 2002

Layer	**NH_4**	**NO_3**	**N_{TOT}**	**O_2**	**PO_4**	**P_{TOT}**	**S**	**SiO_4**	**T**
correlation coefficient									
1	0.51	0.63	0.64	0.80	0.54	0.65	0.58	0.32	0.97
2	-0.08	0.64	0.54	0.76	0.43	0.38	0.85	0.32	0.81
3	0.37	0.18	0.58	0.72	0.56	0.52	0.93	0.59	0.55
bias									
1	1.3	1.59	1.07	0.97	1.56	1.27	0.94	0.76	0.76
2	1.21	1.43	1.05	0.96	1.22	1.13	0.95	0.84	1.02
3	0.62	1.08	1.03	1.28	0.80	0.73	0.86	0.64	1.05
effectiveness									
1	-0.06	0.15	0.16	0.61	-0.09	0.24	0.03	0.05	0.94
2	-0.52	0.19	0.26	0.52	0.05	0.08	0.64	-0.04	0.65
3	0.05	-0.17	0.33	0.41	0.16	0.05	0.46	-0.23	0.05

Table 6 Coefficients of correlation, biases and effectiveness for state variables at the areas of the Gulf of Gdansk (area 1 and 2) and the southern Baltic (area 3) in 1994 – 2002

Area	**NH_4**	**NO_3**	**N_{TOT}**	**O_2**	**PO_4**	**P_{TOT}**	**S**	**SiO_4**	**T**
correlation coefficient									
1	0.47	0.58	0.44	0.74	0.43	0.42	0.50	0.28	0.97
2	0.49	0.67	0.60	0.91	0.85	0.83	0.92	0.88	0.94
3	0.28	0.75	0.49	0.89	0.68	0.70	0.95	0.74	0.95
bias									
1	1.37	1.66	1.35	0.97	1.53	1.21	0.93	0.62	0.98
2	0.67	1.25	1.04	1.03	1.04	1.03	0.92	0.80	0.99
3	1.25	1.19	0.92	0.97	0.94	0.85	0.94	0.74	0.97
effectiveness									
1	-0.12	0.07	-0.21	0.53	-0.18	-0.03	0.03	-0.02	0.93
2	0.14	0.34	0.34	0.76	0.71	0.69	0.61	0.65	0.89

3	-0.12	0.52	0.17	0.78	0.36	0.28	0.78	0.25	0.90

5. Statistic characteristics of the model quality

Defined correlations demonstrated the strength of relationships between the modelled and the measured values, irrespective of whether the modelled simulations were overstated or underrated in relation to the measured ones. Absolute bias testifies to the relationship of the modelled values and observations (Fig. 6). The modelled simulations were overstated for nitrogen compounds at most stations, while underrated for phosphorus compounds. Contents of silicate silicon were definitely underrated at all station. Dissolved oxygen simulations were slightly overstated in the Gdansk Deep (station P1) and little underrated than observed ones at the others (station: K, ZN2, NP). Modelled temperature values diverged the least from the observed ones, while salinity simulations were relatively little underrated in relation to the measured ones. In comparison with simulations nitrate nitrogen only at the shallow station K was underestimated and at the others overstated. Ammonium and total nitrogen concentration measurements were underrated at the stations P1 and K (like at the all ones together), while at the others located near the mouth of the Vistula River (ZN2 and NP) were overstated (Fig. 6).

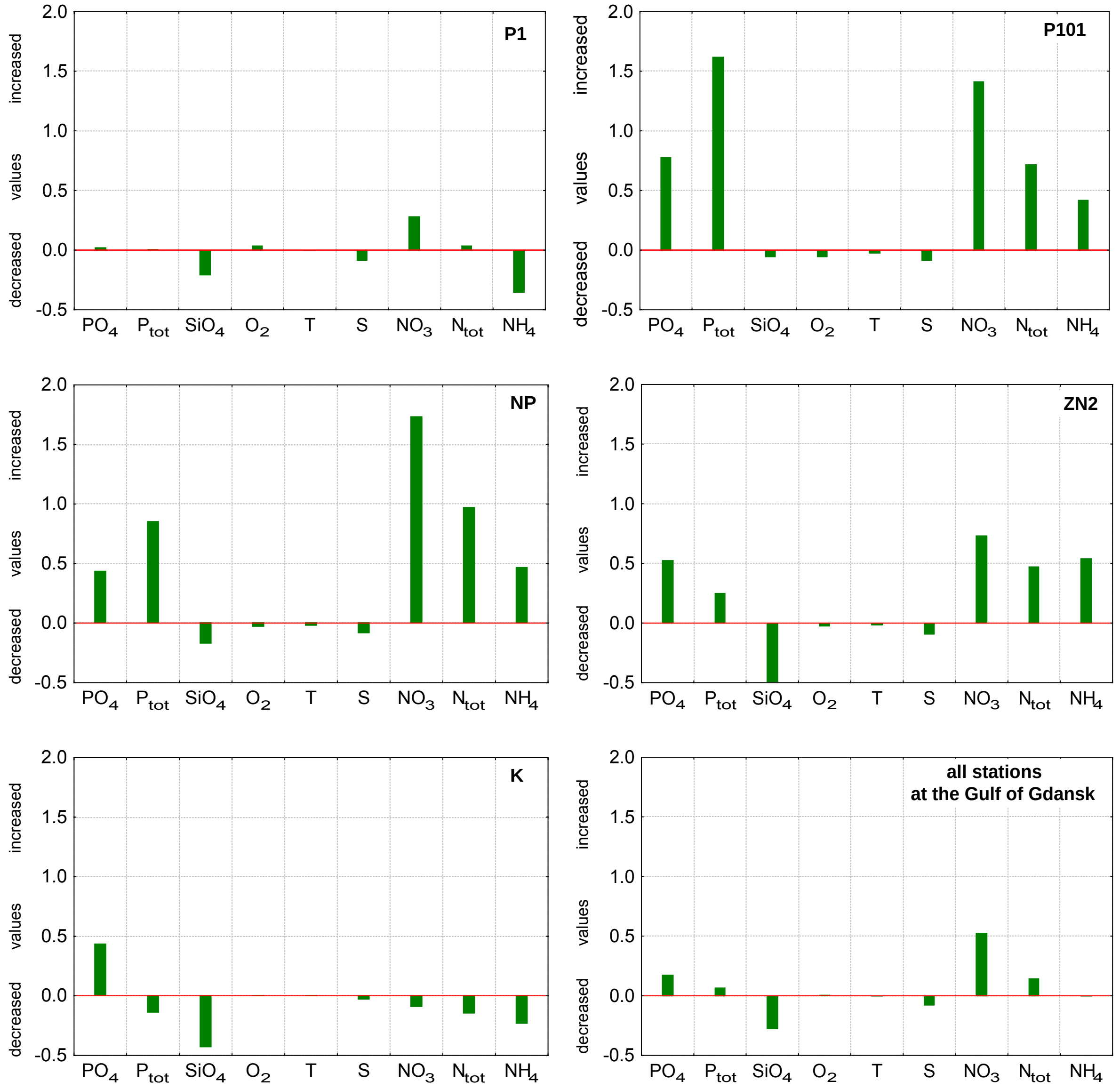


Fig. 6 Absolute bias of the model calculated for state variables of ProDeMo model at stations P1 P101 NP ZN2 K and all stations for 1994–2002

The smallest range of divergence between modeled and measured variables of state appeared at the intermediary layer. Simulation of the temperature, salinity and dissolved oxygen were very close to observations there nearly in the every layer and areas. Modelled silicates silicon were underestimated all over the places but phosphates only in the bottom layer whereas nitrogen and phosphorus compounds were overestimated, particularly phosphates were overstated more than 50 % at the surface layer as well as at the area in front of the mouth of the Vistula river (Fig. 7).

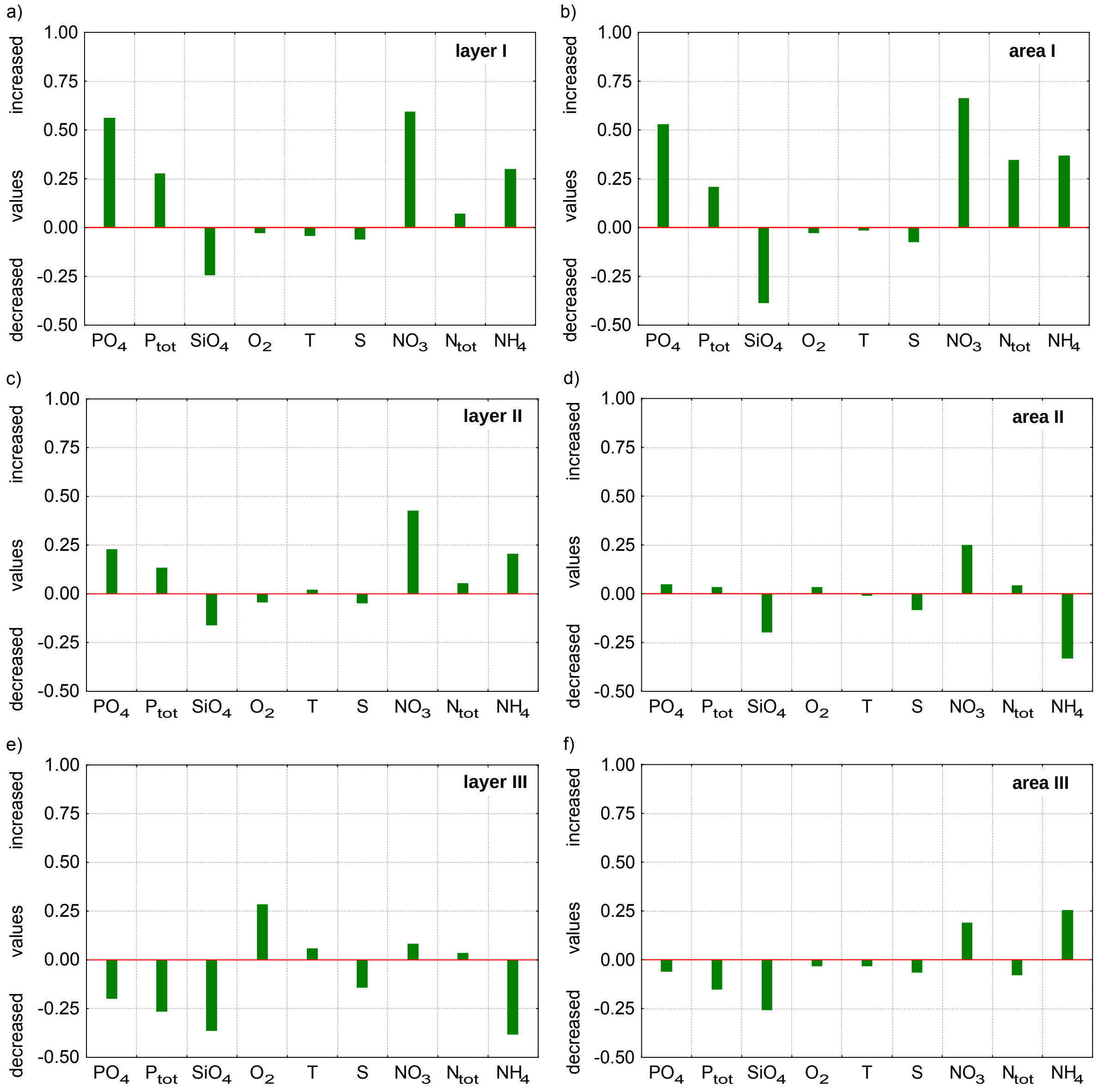


Fig. 7 Absolute bias of the model calculated for state variables of ProDeMo model at the surface intermediate and bottom layers as well as areas placed at the Gulf of Gdansk in 1994–2002

Correlation coefficients combined with such supplementary criteria as conditional and unconditional bias and gave an idea of correlation effectiveness (Fig. 8). It means effective improvement of simulation with respect to observations. The correlation effectiveness expressed by determination coefficient which equals R^2 is often an overstated value, as it reveals the strength of relationship when there is no model bias. By reducing effectiveness Nash-Sutcliff coefficient takes into account the bias and expresses real correlation of compared values. In the case of heavy bias, this effectiveness declines to zero or even reaches negative values indicating no effectiveness at all (Fig. 8a - f).

Effectiveness of the simulations were referred to layers and areas. There have been assigned surface, intermediate and bottom layers as well as three areas: near mouth of the Vistula river, open boundary of the Gulf of Gdansk and southern Baltic. The variables at the open part of Gulf of Gdansk were simulated the most effectively. Physical parameters like temperature, salinity and oxygen dissolved have reached effectiveness coefficients in range 0.62 – 0.9. Silicon and phosphorus compounds were demonstrated with effectiveness of 0.68 – 0.71 but nitrogen ones of 0.05 – 0.47. Simulations of variable states at the southern Baltic waters have got even high (0.78 – 0.9) positive coefficients, for physical parameters. At the area close to mouth of the Vistula river obtained estimation of effectiveness for the temperature and oxygen dissolved was pretty good but for remaining parameters rather weak even for phosphates and silicon negative.

The modeled vertically structure of the state variables were differentiated. The highest values of effectiveness were gained for the temperature and oxygen but opposite for phosphates and silicates at the surface layer (Fig. 8a,c,e). At the intermediate layer only effectiveness for nitrogen ammonia was negative, the other values were satisfying. At the bottom layer obtained coefficients were beneath 50% of effectiveness.

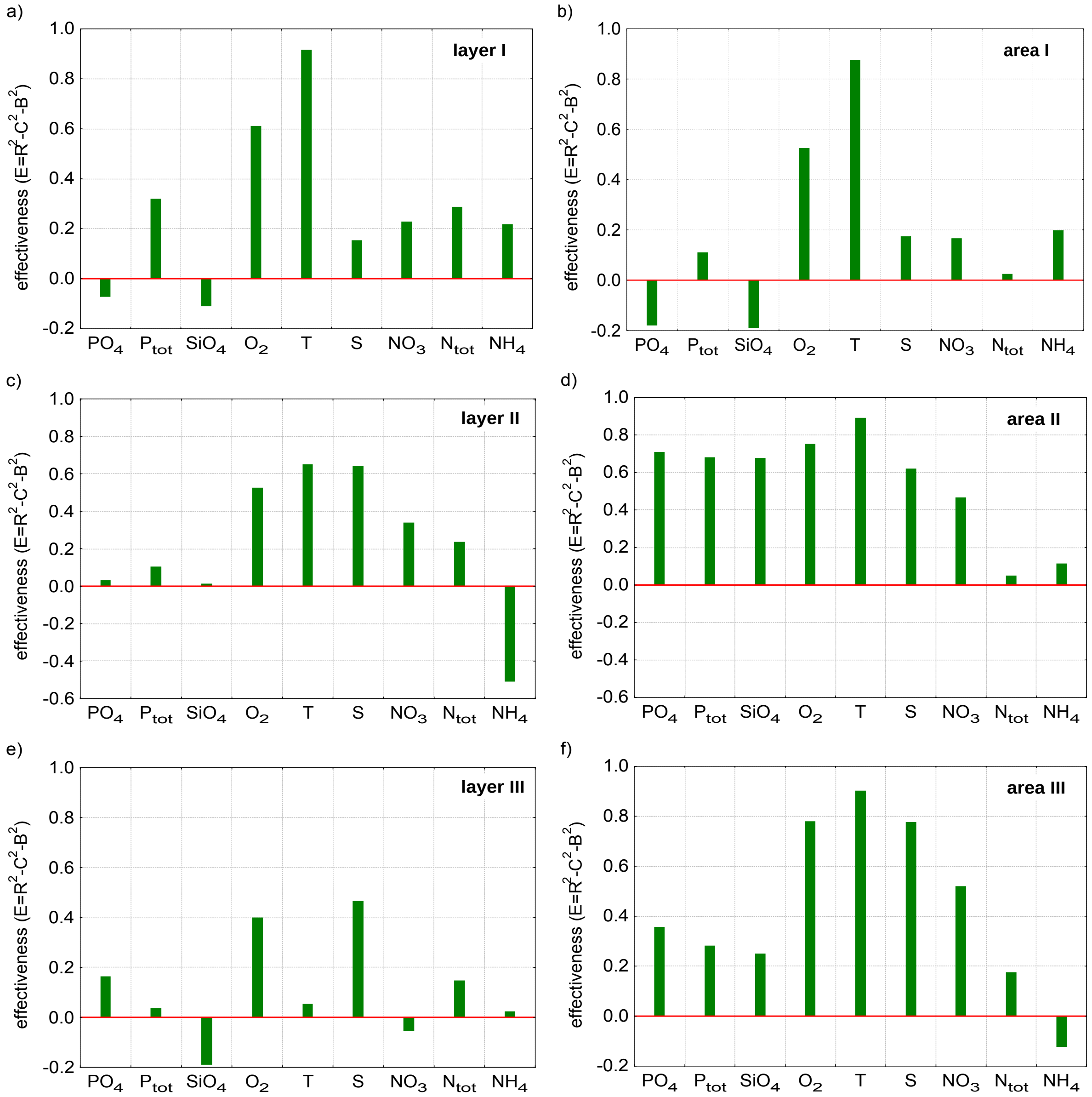


Fig. 8 Coefficients of effectiveness for state variables at layers: a) surface, c) intermediate, e) bottom and areas: b) near the mouth of the Vistula river, d) open boundary Gulf of Gdansk, f) Baltic southern in 1994 – 2002

With all observation stations considered, the best effectiveness was achieved for temperature, salinity and oxygen dissolved in water, which confirms good quality of the hydrodynamic model. The best simulated phosphorus and silicate compounds reached even over 65% and nitrates more than 50% effectively. The estimation of effectiveness largely declined for silicate silicon, total nitrogen and nitrate nitrogen, while they were even negative for silicate silicon and ammonium nitrogen (Fig. 8).

Comparing simulations with observations by means of a special correlation coefficient in the function of total quadratic error a statistical analysis was also made (Fig. 9, 10). Physical parameters were excellently correlated with regard to surface layer (Fig. 9a). Well simulated were phosphorus, silicon and nitrogen compounds except ammonium nitrogen. Complete profiles of two stations extreme located at the Gulf of Gdansk P1 and ZN2 showed very good correlation for temperature, salinity, oxygen dissolved also nitrate and silicon at P1, good for phosphorus and nitrogen compounds at station P1, phosphorus at ZN2 and K as well as satisfied for silicate and nitrogen at ZN2. These results appeared as evident influence of the Vistula river plume (Fig. 9b, c).

All measurements from all stations and times were used for particular variables. The best correlated was salinity and total nitrogen, the weakest – ammonium and nitrate nitrogen (Fig. 9d). The parameters best simulated by the model were: salinity, water temperature and oxygen, well simulated were total phosphorus and nitrogen and phosphate phosphorus. Silicate silicon, nitrate nitrogen and ammonium nitrogen simulations clearly fall behind the others.

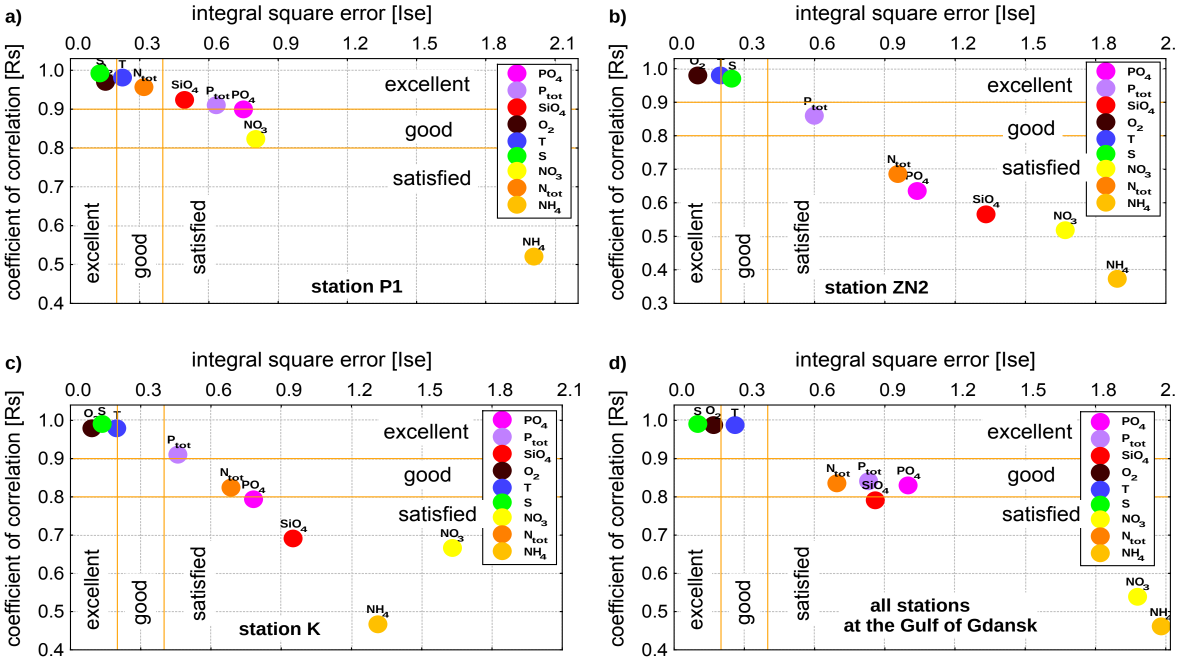


Fig. 9 Special coefficients of correlation in the function of total quadratic error for whole profiles state variables at a) the station P1 b) station ZN2 c) station K d) all stations and terms from the period 1994-2002

Taking into account all measurements from all stations and times used for particular variables the correlation was considered for vertical (layers) and horizontal (areas) their structure. The surface layer parameters are located somewhat lower on the graph (Fig. 10e) than intermediary and bottom layer ones (Fig. 10a, c, e) thus indicating that the simulations departed more from the measured values. In this layer the best correlation between simulations and measurements was obtained by temperature, salinity and oxygen dissolved, good – total phosphorus and nitrogen but remaining parameters only satisfying.

Considering the variables spatially according to areas evaluated at waters in front of the Vistula river showed, temperature, salinity, oxygen dissolved were simulated excellently, total phosphorus quite well, remaining as satisfied (Fig. 10b). At the area of open part the Gulf of Gdansk most modeled variables reached the best correlation with the observations. The similar situation appeared at the intermediary layer. At the waters of the southern Baltic the simulations correlated with measurements getting coefficients decreasing along the diagonal line, from the best value for salinity to the least adequately for ammonium nitrogen (Fig. 10d, f).

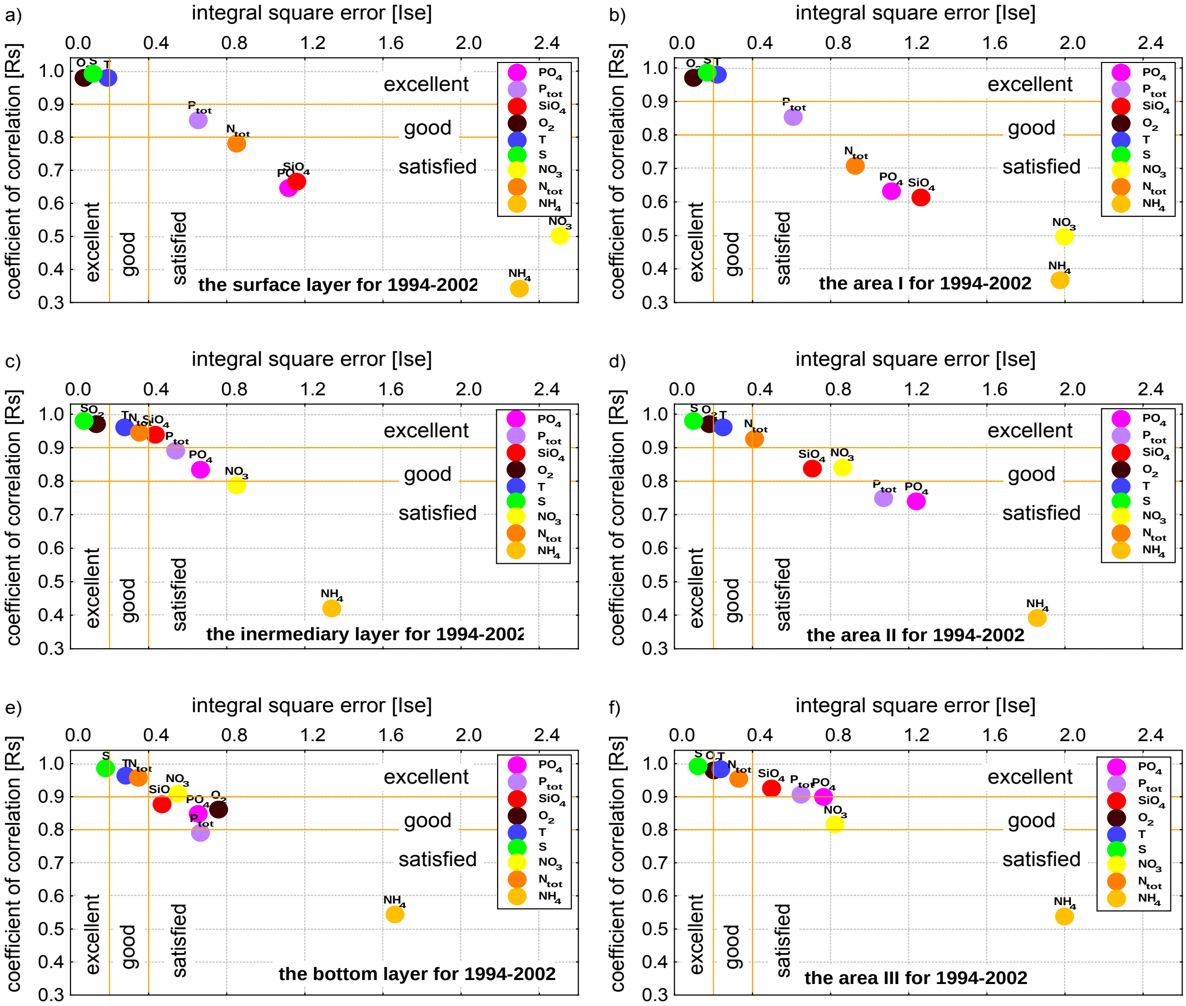


Fig. 10 Special coefficients of correlation in the function of total quadratic error for state variables for all measurements at (left panel): a) surface layer c) intermediate layer e) bottom layer and (right panel) b) area in front of the Vistula river d) open boundary Gulf of Gdansk f) southern Baltic from the period 1994-2002

The statistical measures describing the quality of the model and based on mean quadratic error (correlation and bias coefficients) and total quadratic error (special correlation coefficient), confirm a very good quality of the simulations modelled for physical parameters, good for phosphate phosphorus and silicate silicon and poorer for nitrogen compounds.

CONCLUSIONS

Drawing a conclusion from the former calibration of the model and using the statistics measures, it has been elaborated new favourable calibration. That approach resulted in new set of coefficients to be completed the equations biogeochemical processes. The outcome obtained from comparison between model simulations and measured data as high coefficients of effectiveness gave evidence that the processes have been described accurately as well as the model has got better quality. The model was run during annual cycle, repeated for three years, nine and fifteen years periods. The model replied seasonal changes of nutrient concentrations and phytoplankton biomass. It has appeared as long term stable and complied with the mass conservation law, what certified its internal logic and correctness. The model was subjected to validation using measurements taken from coast waters in front of the mouth of the Vistula river. The obtained compliance with the observations confirmed that the model was a tool for reliable prediction of ecosystem behaviour in the southern Baltic.

Acknowledgements:

The model was developed during the project EUROCAT No EVK1/2000/00510 founded by EU.

Literature:

Kannen A., Jedrasik J., Kowalewski M., Oldakowski B., Nowacki J., 2004, *Assessing catchment-coast interactions for the Bay of Gdansk*, in: G. Schernewski & N. Löser (eds.): Managing the Baltic Sea, Coastline Reports 2 (2004), 155 - 165

Kowalewski, M., J. Jedrasik, B. Oldakowski & J. Nowacki (2003): *The impact of the Vistula river on the coastal waters of the Gulf of Gdansk*, 1 - 70

Ołdakowski B., Kowalewski M., Jędrasik J., 1994, *The nutrient dynamic model for the Gulf of Gdańsk*, Proc. Conf. Baltic Oceanogr., Sopot, Poland, 528-543

Ołdakowski B., Renk H., 1997, The *conception and structure of the Production-Destruction of Organic Matter Model; verification tests for the Gulf of Gdańsk*, Oceanol. Stud., 26 (4), 99-122

Węglarczyk S., 1998, *The interdependence and applicability of some statistical quality measures for hydrological models*, J. Hydro., 206, 98-103

Appendix I Statistical measures

The basic compared quantities were the modeled values – y and the observed values – x. The differences between them were defined as **a model error**:

$$\Delta xy = y - x \tag{1}$$

The **mean quadratic error** was a measure based on error quantity:

$$E_{rs} = \overline{(\Delta xy)^2} \tag{2}$$

The **absolute bias** of the model was expressed by difference of averages:

$$Q_m = \overline{\Delta xy} = \bar{y} - \bar{x} \tag{3a}$$

It is a dimensional quantity showing to what degree the modelled values were underestimated or overestimated in relation to the observed ones.

$$Q_{mo} = \frac{\bar{x}}{\bar{y}} \tag{3b}$$

The correlation coefficient is a product of calculated and observed standard quantities:

$$r = \frac{\overline{(x-\bar{x})\cdot(x-\bar{y})}}{S_x \cdot S_y} = \frac{\operatorname{cov}(x,y)}{S_x \cdot S_y} = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{S_x \cdot S_y} \tag{4}$$

where: x– the observed value of state variable; y – the modelled value of state variable; $\bar{x}$ – mean value of the observed state variable; $\bar{y}$ – mean value of the modeled state variable; cov(x,y) – co-variance of the observed and modeled values; standard deviation of the observed and modeled values;

$$S_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{N}} \quad S_y = \sqrt{\frac{\Sigma(y-\bar{y})^2}{N}};$$

The coefficient of correlation is not “sensitive” to the characteristics known as model bias (Węglarczyk, 1998). Relationship between correlation coefficient and bias is contained in the mean quadratic error E_{rs}

$$E_{rs} = S_x^2\left[(1-r^2) + \left(\frac{S_y}{S_x} - r\right)^2 + \frac{\bar{Q_m}^2}{S_x^2}\right] \tag{5}$$

The second term of the equation (5) describing the correlation between the model error and its simulation will be defined as **conditional bias** and denoted as C^2

$$C^2 = \left(\frac{S_y}{S_x} - r\right)^2 \tag{6}$$

The third term of equation (5) expresses **unconditional bias** B^2 defined as the ratio of absolute bias to standard deviation of observations:

$$B^2 = \frac{\bar{Q_m}^2}{S_x^2} \tag{7}$$

Nash and Sutcliff (Węglarczyk, 1998) divided (5) by S_x^2 keeping C^2 and B^2 and assuming that $1 - \frac{E_{rs}}{S_x^2} = E$ obtained the **coefficient of effectiveness E** in form:

$$E = r^2 - C^2 - B^2 \tag{8}$$

After removing of bias it expresses correlation effectively. This coefficient was assumed as basic for optimization of model calibration. If there is no bias, then it equals to square coefficient of correlation and called as coefficient of determination $\boldsymbol{E_d}$.

$$E_d = r^2 \tag{9}$$

The relation of correlation coefficient with the **total quadratic error**

$$E_{rc} = \frac{\sqrt{E_{rs}}}{\bar{x}} \tag{10}$$

leads, to correlation of the so called **special coefficient of correlation *Rs*** in relation to ***Erc*:**

$$R_s = \sqrt{1 - \frac{E_{rs}}{S_x^2 + \bar{x}^2}} \tag{11}$$

These formulas were assumed as an algorithm for estimation of the simulations by using calibrating database.