Semerádová, S., Černíková, A.: 2026. Relationship between hydrological situation and water quality. In: Tolasz, R., Šustková, V. (Ed.), Proceedings of the 2nd PERUN Project Conference (TA ČR, SS02030040). CHMI, Praha, s. 146–153, ISBN 978-80-7653-095-9. https://doi.org/10.59984/978-80-7653-095-9.18

Relationship between hydrological situation and water quality

(Vztah hydrologické situace a jakosti vod)

Semerádová, S., Černíková, A.

Silvie Semerádová, T. G. Masaryk Water Research Institute, silvie.semeradova@vuv.cz

Alena Černíková, Department of Mathematics, Faculty of Science, Jan Evangelista Purkyně University Ústí Nad Labem, alena.cernikova@ujep.cz

https://doi.org/10.59984/978-80-7653-095-9.18

Abstract: Climate change is altering the hydrological regime, which affects both the availability and quality of water resources. This study focuses on a statistical evaluation of the relationship between flow and the concentration of less frequently monitored pollutants of various origins: fluoranthene (PAHs from atmospheric pollution), metabolites of the pesticides acetochlor and metolachlor (agricultural pollution), and AMPA (a metabolite of glyphosate, from agricultural and other sources of pollution). The analysis utilized extensive panel data from operational monitoring conducted between 2010 and 2019 at 148 monitoring sites in the Czech Republic, paired with daily flow data from the Czech Hydrometeorological Institute (CHMI). The data were analyzed using the R software package via the Generalized Additive Model (GAM), which accounted for nonlinear temporal trends, seasonality, spatial dependence on flow, and autocorrelation structure. The methodology used has proven to be an effective tool that, with appropriate modifications, could be used to predict changes in surface water quality, optimize monitoring networks, and serve as a basis for watershed quality models.

Keywords: Surface water quality – Hydrological conditions – GAM model – fluoranthene – pesticides

Abstrakt: Změna klimatu mění hydrologický režim, což ovlivňuje dostupnost i kvalitu vodních zdrojů. Tato studie se zaměřuje na statistické vyhodnocení vztahu mezi průtokem a koncentrací méně často sledovaných polutantů s odlišným původem: fluoranthenu (PAU z atmosférické zátěže), metabolitů pesticidů acetochloru a metolachloru (zemědělské znečištění) a látky AMPA (metabolit glyfosátu, ze zemědělského i dalšího znečištění). Na několika povodích různého typu jsou ukázány proměny vybraných parametrů jakosti povrchových vod v závislosti na období a hydrologické situaci. Pro analýzu byla využita rozsáhlá panelová data provozního monitoringu z let 2010–2019 ze 148 profilů v ČR, spárovaná s denními průtoky z ČHMÚ. Data byla vyhodnocena v softwaru R pomocí Generalizovaného Aditivního Modelu (GAM), který zohlednil nelineární časový trend, sezónnost, prostorovou závislost na toku a autokorelační strukturu. Použitá metodika se ukázala jako efektivní nástroj, který by při vhodné úpravě bylo možné použít pro predikci vývoje jakosti povrchových vod, optimalizaci monitorovacích sítí i jako podklad pro jakostní modely povodí.

Klíčová slova: jakost povrchových vod – hydrologická situace – GAM model – fluoranthen – pesticidy

Introduction

As the climate changes, the hydrological conditions also change, which affects not only the availability of water resources but also, potentially, their quality. To predict trends in surface water quality parameters, it is necessary to examine the relationships between hydrological conditions and water quality.

A dataset was selected and statistical models were tested for a detailed statistical analysis. Given that the Czech Republic has both long-term flow monitoring – conducted at selected stations by the Czech Hydrometeorological Institute (CHMI) and operational monitoring of surface waters, carried out regularly by river basin authorities, which measure a wide range of substances at hundreds of monitoring sites every month, the data set is sufficient; the key is to prepare the data correctly and select an appropriate statistical model.

The results of such an assessment can be used to predict quality trends, but also, for example, to optimize monitoring campaigns or as a method in watershed quality models.

Similar statistical analyses yield very good results for nutrients. These have been addressed from various perspectives in other studies (Minaudo 2019; Rosendorf 2016).

Therefore, for the purposes of this study, we decided to focus on less frequently assessed indicators – fluoranthene, for which a new limit for good status has been established as a representative of the group of polycyclic aromatic hydrocarbons under the proposed amendment to Directive 2008/105/EC of the European Parliament and of the Council (tightening the limit from the current $6.3 \cdot 10^{-3}$ $\mu g \cdot l^{-1}$ to $7.62 \cdot 10^{-4}$ $\mu g \cdot l^{-1}$). PAU compounds enter the aquatic environment primarily through the atmosphere as combustion products. Given the improving air quality, this represents both a current and a long-standing source of pollution (Sýkora 2026).

The second group of substances, on the other hand, consists of "new" pollutants, specifically, certain pesticides and their metabolites, whose monitoring has become more widespread with the development of analytical methods (Kodeš 2023).

Data

To compare quality and hydrological data, a set of profiles and operational monitoring indicators was selected, along with the corresponding daily discharge values recalculated based on the nearest hydrological station with direct measurements. The criteria for data selection were that measurements should be available for the broadest possible range of indicators (with particular emphasis on indicators causing failure to achieve good status according to the Water Framework Directive), there was a hydrological station nearby for converting flow rates, and there were sufficient data points above the limit of quantification.

Operational monitoring data from 2010–2019 were used as the quality data. Flow measurement profiles were assigned to quality profiles based on geographic and hydrological proximity – the stations selected were those closest to the quality profile, either downstream or upstream. The actual flow rates were scaled based on the ratio of catchment area. Profiles with a catchment area ratio greater than 4 or less than 0.2 were excluded.

Daily discharge data from hydrological stations with continuous measurements were used to assess the hydrological situation. Based on a long-term series of daily discharge data from 1980 to 2019 (in a few cases, only a shorter series was available, but at least from 2002 to 2019), discharge percentiles were determined for each profile. The recalculated flow rate at the time of sampling for quality determination was compared with the (recalculated) percentile and classified into the corresponding group. This made it possible to determine whether the flow rate was significantly increased or decreased at the time the samples were taken.

Method

First, the measurability of selected indicators was evaluated. The number of samples with concentrations below the limit of quantification was compared with the hydrological conditions.

The data were then analysed using hydrological statistical models. In these models, concentrations below limit of quantification (LOQ) were replaced with the lowest measurable value. A drawback of

this approach is that the limit of detectability could vary across different stations for the same substance. However, omitting these values would mean losing information about the best results, which would skew the conclusions even further.

Panel data with monthly time series from 2009 to 2019 for 148 monitoring stations were thus available for the statistical models. For each variable, the long-term trend, seasonality, and dependence on flow were determined. Flow was represented here by a categorized percentile for each monitoring station. This flow percentile was included in the model as an ordered factor. The data were analysed using an additive GAM model (Generalized Additive Model) (Wood and Augustin 2002; Wood 2017) for panel hydrological data, which allows spatial dependence to be incorporated into the model (values downstream are influenced by values at upstream monitoring stations).

Long-term trends were modelled nonlinearly using a smoothing spline with a maximum basis dimension of 5. Seasonality was assessed as a function of a categorical variable with 12 categories (individual months), and the dependence on flow was calculated as a polynomial of higher degree. In addition to the main dependencies under investigation, a linear dependency on data from the previous measuring station was also included in the model. The correlation structure of the data was incorporated into the model using a random effect for the measuring station, which affects the absolute term, and an AR(1) dependency on the value from the same measuring station in the previous month.

Since the data contained outlier values for all variables, the logarithmic transformation of the original values were used in the analyses.

The data were processed in Excel and in R software, version 4.5.1, within the RStudio environment using the mgcv library.

By combining data from ten years of measurements across 148 measuring stations, a large dataset was obtained in which all tested relationships were found to be statistically significant at a high level of significance. For this reason, statistical significance is not commented on in the results. The conclusions are interpreted primarily using graphical outputs depicting the estimated relationships and specific estimates of the regression coefficients.

Discussion

The analysis of fluoranthene utilized 10,399 data points obtained from all 148 monitoring stations across all years studied. In 2009, data were available from only one monitoring station, namely Moldava in Varnsdorf. Since 2010, data have been available from at least 80 stations each year. Individual watercourses vary significantly in their fluoranthene content; the highest concentrations are found in the Zákolanský Stream, the Lučina, the Oder, and the Hvozdnice, where the average concentration exceeds 0.05 µg/L. The summary results for fluoranthene show a slight decline over time with irregular seasonal fluctuations. The highest concentrations tend to occur during the summer months, with average values for June and July exceeding 0.02 $\mu g \cdot l^{-1}$, while in other months they reach a maximum of around 0.015 $\mu g \cdot l^{-1}$. A visible peak in the estimated model corresponds to the floods in June 2013 (Fig. 1). Extreme concentrations exceeding 0.4 were measured that month on the Opava and Oder rivers.

For fluoranthene, the flow-rate dependence is dominant: at low to moderate flow rates, concentrations remain constant and increase exponentially as flow rates rise (Fig. 2). A similar curve is observed when comparing the number of measurements below the limit of quantification. In total, there were 376 values below the limit of detection in the dataset, of which only 8 occurred at high flow rates in the 90–95th and 95^{th} –100^{th} percentiles.

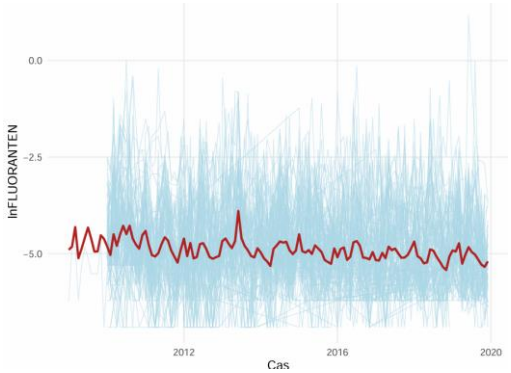


Fig. 1. Time-series of the logarithm of flouranthene. The blue curves show the time-series at individual monitoring stations; the red curve shows the average trend estimated by the hydrological model.

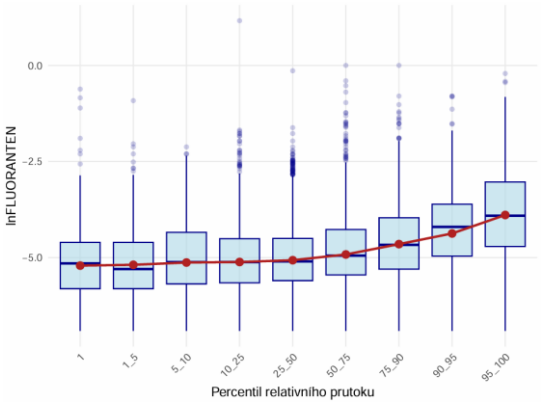


Fig. 2. The dependence of the fluoranthene logarithm on the relative flow percentile. For each percentile, a box plot is shown, indicating the minimum and maximum non-outlier values, the quartiles, the median, and the outliers. The red line corresponds to the mean values estimated by the GAM model.

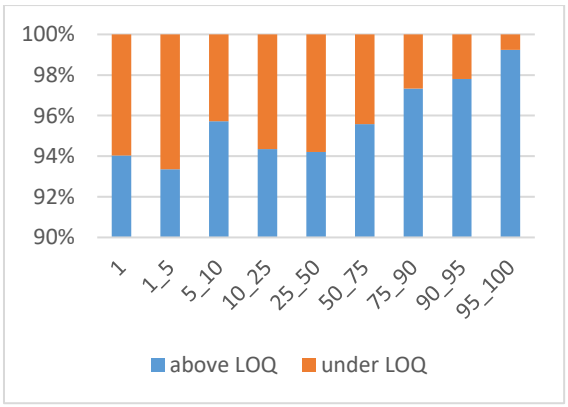


Fig. 3. The relationship between the number of values below the detection limit and the percentile of relative flow.

When it comes to pesticides, the situation is more complex in terms of both data availability and comparability. Metabolites of the pesticide acetochlor were selected as an example. A total of 9,920 observations from all 148 monitoring stations and from the years 2010 through 2019 were available for analysis. The use of acetochlor was banned in the EU in 2012. As a result, it can be expected that the trend in the concentration of its metabolites in water will not remain constant throughout the entire monitoring period. For this reason, we modified the model to allow for changes in seasonality over the years. The estimated curve showing the trend of acetochlor.mb over time indicates that an increase can still be observed even after 2012. A turning point occurs in 2015, after which a slight decline becomes apparent. Following the ban and the cessation of use, the previously clear seasonal pattern which was

characterized by high concentrations primarily at the beginning of summer, specifically in May and June, when average concentrations ranged from 0.08 to 0.09 $\mu g \cdot l^{-1}$, while in other months average concentrations were approximately 0.06 $\mu g \cdot l^{-1}$ has disappeared (these averages also include years in which the seasonal pattern had already disappeared), Fig. 4.

As the flow rate increases, the variability in concentrations increases as well; particularly at the beginning of the monitoring period, high flow rates during the spring months may have caused very high concentrations of this substance.

For comparison, the trend in the occurrence of metabolites of the pesticide metolachlor is shown (it has also been banned in the EU, but not until 2024, so this is not reflected in these data). A total of 10,176 data points were available. Here, too, an increase in the occurrence of these metabolites is evident until 2015, when their levels in watercourses leveled off. Unlike acetochlor metabolites, however, a noticeable seasonal pattern persists throughout the entire monitoring period, although it also weakens slightly over time (Fig. 5). As with the previous substance, the highest concentrations occur in May and June, at approximately 0.1 $\mu g \cdot l^{-1}$, compared with an average concentration of 0.6 $\mu g \cdot l^{-1}$ in the other months. The dependence on flow rate is also the same as for the previous substance – concentration variability increases.

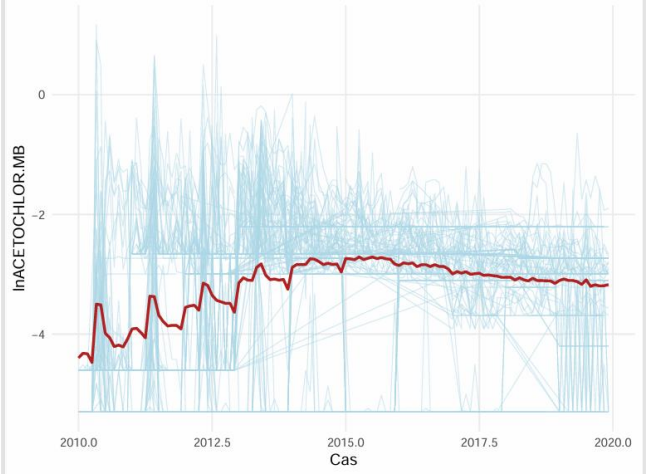


Fig. 4 Time-series of the logarithm of acetochlor metabolites. The blue curves show the time-series at individual monitoring stations; the red curve shows the average trend estimated by the hydrological model.

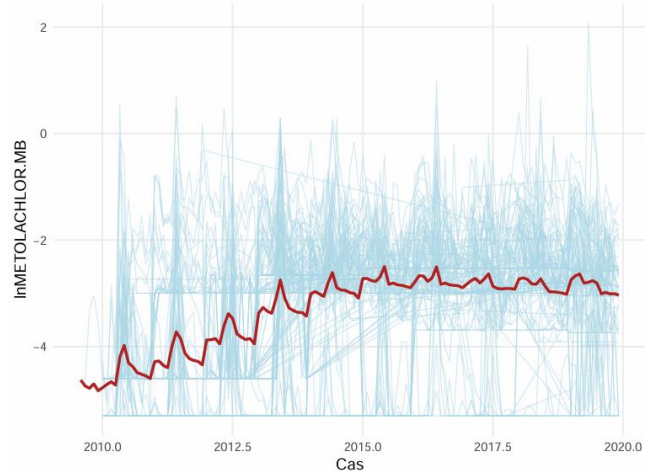


Fig. 5. Time-series of the logarithm of metolachlor metabolites. The blue curves show the time-series at individual monitoring stations; the red curve shows the average trend estimated by the hydrological model

For the sake of interest, AMPA (aminomethylphosphonic acid) the main degradation product (metabolite) of glyphosate, the world's most widely used herbicide has been included. For this substance, we had only 7,050 data points available from 2010 to 2019, but there were significantly fewer data points in the first two years of the study. The data again come from all 148 monitoring stations. By far the highest concentrations were measured on the Ostravice River, with an average of 8.8 $\mu g \cdot l^{-1}$; averages exceeding 5 $\mu g \cdot l^{-1}$ were also measured on the Lučina, Hvozdnice, and Lubina rivers, as well as on the Podolský and Černý streams.

Most stations (70) do not reach an average of 1 $\mu g \cdot l^{-1}$. Here, a slight increase over time is evident throughout the entire monitoring period, along with pronounced seasonality – Fig. 6. The highest concentrations are measured in July and August (average values in these months exceed 1.27 $\mu g \cdot l^{-1}$; in September, they exceed 1.2 $\mu g \cdot l^{-1}$), compared to values around one in the other months. Furthermore, a decreasing dependence on flow rate is evident; this decrease is observed between the 10th and 90th percentiles of the flow rate; under extreme water conditions, there is no significant change in concentration (Fig. 7). Similarly, we see an increase in the number of values below the limit of determination at higher flow rates (Fig. 8).

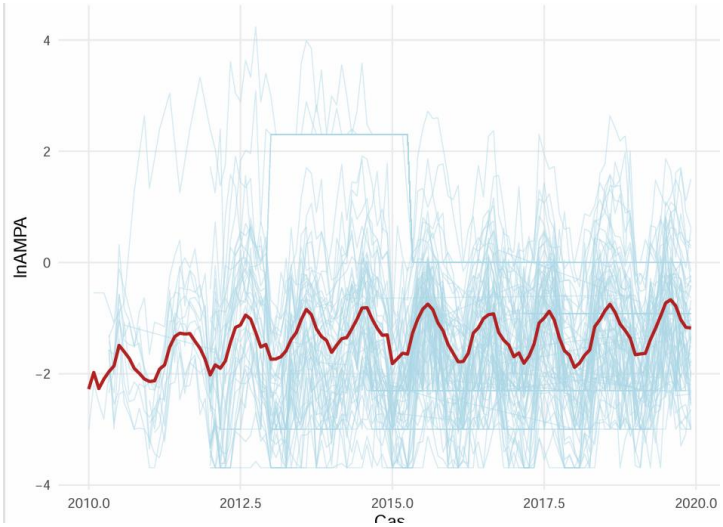


Fig. 6. Time-series of the logarithm of AMPA. The blue curves show the time-series at individual monitoring stations; the red curve shows the average trend estimated by the hydrological model.

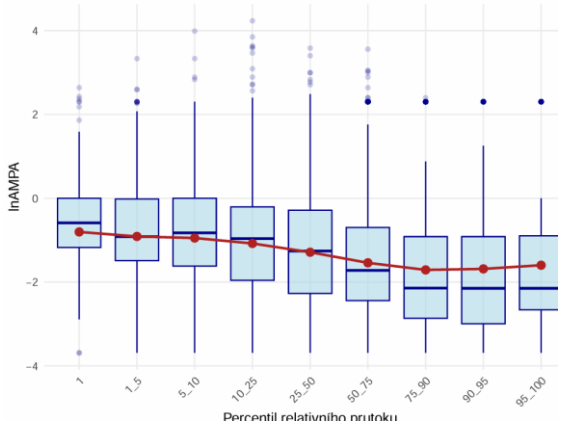


Fig. 7. The dependence of the AMPA logarithm on the relative flow percentile. For each percentile, a box plot is shown, indicating the minimum and maximum non-outlier values, the quartiles, the median, and the outliers. The red line corresponds to the mean values estimated by the GAM model.

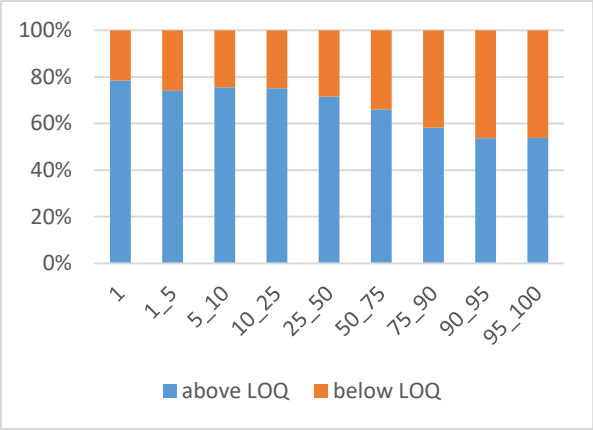


Fig. 8. The relationship between the number of values below the detection limit and the percentile of relative flow.

Conclusion

In this paper, we focused on three types of substances: fluoranthene, for which flow-rate dependence is the dominant factor; pesticide metabolites, for which a time trend dependent on the source of the pesticide itself is evident; and AMPA, for which a dominant seasonal pattern is evident.

It is evident that pesticide metabolites reach a certain saturation level in watercourses that they do not exceed, even with a constant influx of pesticides (see metolachlor). Conversely even if the influx of new pesticides ceases, it will still take many years for the watercourses to be cleared of these metabolites (see acetochlor).

The methods used have proven to be effective, and the relationships have been demonstrated. At the same time, the available datasets provide a sufficiently robust data foundation for evaluating the prevailing relationships and, consequently, for identifying potential sources or predicting future trends. The models described can be applied with even greater reliability to more comprehensive and longer-term nutrient monitoring time series.

Acknowledgments

The research was partially funded by the Technology Agency of the Czech Republic (grant no. SS02030040, PERUN).

We would like to thank the river basin authorities and the Czech Hydrometeorological Institute for providing the data.

References

Kodeš, V., Vejvodová, J., Sirotková, K., 2023. Kvalita povrchových vod z pohledu pesticidních látek a faktory mající vliv na výskyt pesticidů ve vodách. Vodní hospodářství 73(11), pp. 2–8 [online] [cit. 30.6.2026]. ISSN 0042-790X. https://vodnihospodarstvi.cz/kvalita-povrchovych-vod-z%E2%80%AFpohledu-pesticidnich-latek-a%E2%80%AFfaktory-majici-vliv-na-vyskyt-pesticidu-ve-vodach/

Minaudo, C., Dupas, R., Gascuel-Odoux, CH., Roubeix, V., Danis, P. A., et al., 2019. Seasonal and Event-Based Concentration–Discharge Relationships to Identify Catchment Controls on Nutrient Export Regimes. *Advances in Water Resources*, **131**. https://doi.org/10.1016/j.advwatres.2019.103379

Rosendorf, P., Vyskoč, P., Prchalová, H., Fiala, D., 2016. Estimated contribution of selected non-point pollution sources to the phosphorus and nitrogen loads in water bodies of the Vltava river basin. *Soil and Water Research*, **11**, pp. 196–204. https://doi.org/10.17221/15/2015-SWR

Sýkora, F., Semerádová, S., Kodeš, V., Sucharová, J., Verlíková, N., et al., 2026. Atmosférická depozice polycyklických aromatických uhlovodíků v pilotním povodí Výrovky a městských oblastech Prahy a Ostravy. *VTEI*, **68**(1), pp. 22–33. ISSN 0322-8916. https://doi.org/10.46555/VTEI.2025.11.004

Wood, S. N., 2017. Generalized Additive Models: An Introduction with R (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

Wood, S. N., Augustin, N. H., 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling*, **157**, pp. 157–177.