

Využitie družicových dát pre potreby posudzovania kvality ovzdušia

Karol Seják

Anotácia

Využitie družicových dát pre potreby posudzovania kvality ovzdušia bolo demonštrované jednoduchou časovou a priestorovou analýzou počtu dát a hodnôt troposférického stĺpca NO₂ z družice Sentinel-5P. Boli vytvorené mapy s priemernými hodnotami troposférického stĺpca NO₂ za rok 2019 a 2020 pre územie Slovenska. Na základe skúmania viacerých vzťahov medzi družicovými a pozemnými meraniami boli tiež pre dané roky vytvorené mapy prízemných koncentrácií NO₂. Zníženie koncentrácií v roku 2020 oproti 2019 na väčšine územia Slovenska môže byť spôsobené zmenou správania sa spoločnosti súvisiacou s pandemiou COVID-19.

Kľúčové slová: Sentinel-5P, troposférický stĺpec NO₂, prízemné koncentrácie, lineárna regresia

Anotation

The use of satellite data for air quality assessment was demonstrated by a simple temporal and spatial analysis of the number of data and values of the tropospheric NO₂ column from the Sentinel-5P satellite. Maps were created with average values of the tropospheric column NO₂ for the year 2019 and 2020 for the territory of Slovakia. Based on the investigation of several relationships between satellite and ground measurements, maps of ground NO₂ concentrations were also created for the given years. The reduction of concentrations in 2020 compared to 2019 in most of Slovakia can be caused by the change in the behavior of the society related to the COVID-19 pandemic.

Key words: Sentinel-5P, tropospheric NO₂ column, ground concentrations, linear regression

1 Úvod

Copernicus Sentinel-5 Precursor misia je prvou Copernicus misiou zameranou na monitoring atmosféry Zeme. Copernicus Sentinel-5P je výsledkom spolupráce ESA (Európskej vesmírnej agentúry), Európskej komisie, Holandského vesmírneho úradu, priemyslu a vedcov. Misia pozostáva z jedného satelitu nesúceho TROPOsférický monitorovací inštrument TROPOMI, ktorý bol vyvinutý v spolupráci Európskej vesmírnej agentúry a Holandska. Hlavným cieľom Copernicus Sentinel-5P misie je vykonávať atmosférické merania s vysokým časovo-priestorovým rozlíšením pre použitie v problematike kvality ovzdušia, ozónu, UV radiácie a v klimatologickom monitorovaní [1].

Pre potreby monitorovania kvality ovzdušia sa štandardne využíva sieť meračích staníc kvality ovzdušia umiestnená na zemskom povrchu. Lenže inštalácia, údržba a zvyšovanie ich počtu si vyžadujú vysoké investície a náklady [5]. Družicové pozorovania nám môžu poskytovať informácie o kvalite ovzdušia mimo pevnej siete monitorovacích staníc. Monitorovanie kvality ovzdušia je dôležité, pretože vysoké koncentrácie O₃, PM a NO₂ v blízkosti zemského povrchu môžu spôsobiť vážne zdravotné problémy, ako sú pľúčne a kardiovaskulárne choroby a tiež poškodzovať ekosystémy [4].

V tomto príspevku sa zameriame na využiteľnosť Sentinel-5P dát z meraní troposférického stĺpca NO₂ pre potreby monitorovania kvality ovzdušia na Slovensku. Najprv rozanalyzujeme

dostupnosť kvalitných dát z družice, následne budeme detailnejšie analyzovať hodnoty troposférického stĺpca v priestore a čase. Napokon preskúmame vzťah družicových meraní s pozemnými meraniami koncentrácií NO₂, pričom súčasne budeme porovnávať dáta z rokov 2019 a 2020.

2 Troposférický stĺpec NO₂

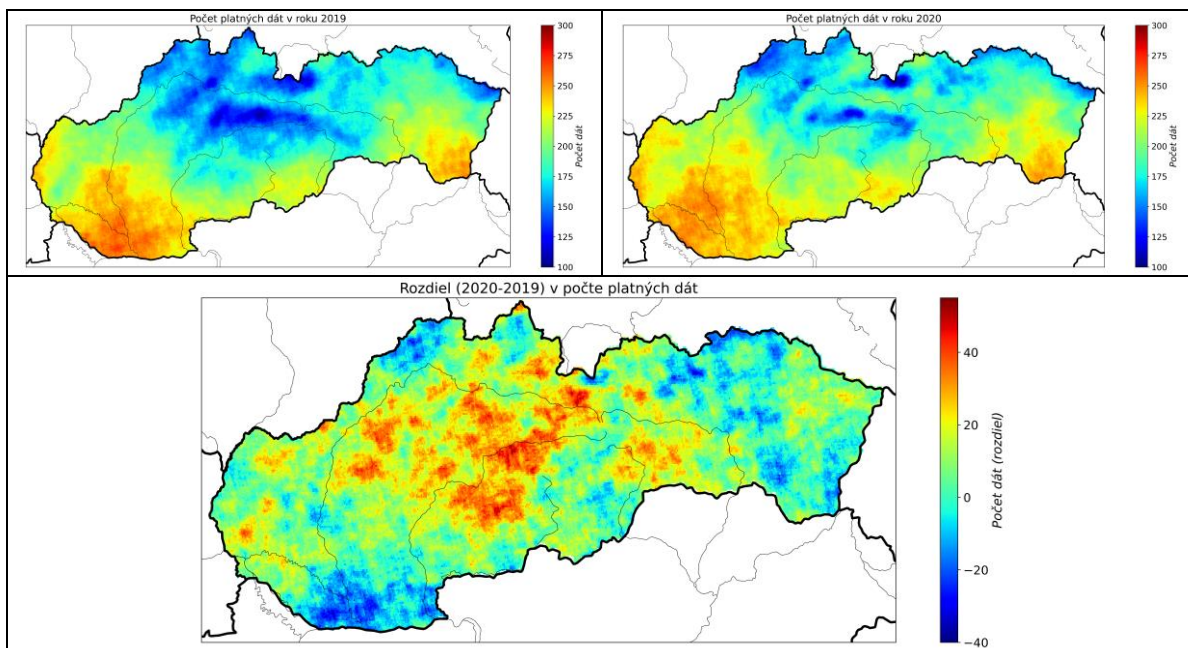
Dáta z družice

Jedným zo základných problémov pri posúdení využiteľnosti produktov z družice Sentinel-5P je dostupnosť kvalitných dát. Priestorové rozlíšenie družice bolo 7 km x 3,5 km, pričom od 6. 8. 2019 je rozlíšenie zvýšené na 5,5 km x 3,5 km [7]. Kvalita údajov závisí od oblačnosti, albeda povrchu, prítomnosti ľadu a snehu. Nie všetky pixely (dáta) sa dajú využiť plnohodnotne. Kvalitu dát charakterizuje hodnota qa , ktorá môže nadobúdať spojité hodnoty od 0 po 1. Hodnota qa predstavuje limitnú hodnotu individuálnej kvality každého pixelu. V prípade troposférického stĺpca NO₂ a celkového stĺpca NO₂ je doporučená hodnota $qa > 0,75$. Nastavením hodnoty $qa > 0,75$ sa vymažú oblasti pokryté oblačnosťou, oblasti pokryté snehom a ľadom a inak problematické údaje [2].

Dáta z družice sa nedajú využiť plnohodnotne hocikde na svete. Najhoršiu pozíciu majú v tomto krajinu, ktoré sú v oblasti vyšších zemepisných šírok, majú veľa oblačnosti a pomerne nízke znečistenie. Príkladom týchto obmedzení je napríklad Nórsko. Nórska štúdia [3], ktorá sa zaoberala priestorovým a časovým rozložením platných údajov v prípade troposférického stĺpca NO₂ v 5 jednotlivých mestských aglomeráciách (doménach) preukázala, že priemerne je platných 20 až 50 % údajov. Časová analýza ukázala, že frakcia platných pixelov v každej doméne má sezónny cyklus dosahujúci maximum 50 až 75 % v letných mesiacoch a 0 až 20 % v zime. Viac platných údajov bolo pri pobreží a menej vo vnútrozemí a v horách kvôli oblačnosti a častiam pokrytým snehom a ľadom. Sezónny cyklus dostupnosti dát je inverzný voči znečisteniu NO₂, lebo maximálne hodnoty NO₂ sa dosahujú práve v zimných mesiacoch. Hodnota kvality dát qa sa pre zenitové uhly vyššie ako 81,2° výrazne znižuje (prenásobenie hodnotou 0,3) a pre zenitové uhly vyššie ako 84,5° je rovná nule. Toto obmedzenie zemepisnou šírkou v praxi znamená, že pre vysoké zemepisné šírky, v ktorých leží aj Nórsko, nie sú okolo zimného slnovratu k dispozícii žiadne dáta, a to aj v prípade, že je jasná obloha [3].

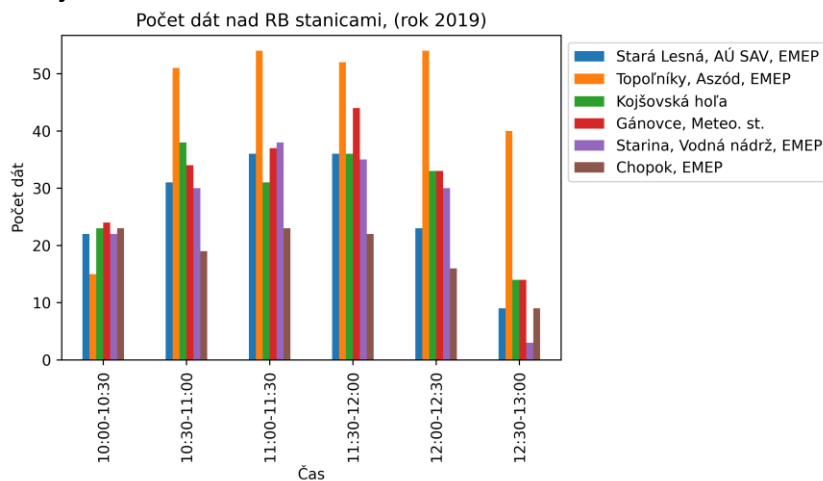
Dáta z družice pre oblasť SR

V prípade územia SR bude hlavným dôvodom nedostatku dát oblačnosť, a to najmä na strednom Slovensku vzhľadom na jeho vysokú nadmorskú výšku. Pre územie SR sme spracovali dáta z rokov 2019 a 2020. Pre zabezpečenie kvalitných dát sme využili filter kvality $qa > 0,75$. Za rok 2019 bol celkový počet družicových snímok nad SR 576 a za rok 2020 to bolo 569 snímok. Počas roku 2019 sa celkový počet platných dát nad daným miestom (mriežkovým bodom) pohyboval v rozsahu od 103 po 272 v závislosti od polohy. V roku 2020 sa tento počet dát pohyboval od 101 po 264. Najmenej platných dát bolo v hornatých oblastiach stredného Slovenska a najviac v oblastiach nížin, tam, kde je počas roka najmenej oblačnosti. Porovnaním platných údajov v rokoch 2019 a 2020 sme zistili mierny nárast platných dát v roku 2020 oproti roku 2019 (Obr.1), a to hlavne v oblastiach stredného Slovenska, avšak viac platných dát možno badať na väčšine územia SR. Z toho vyplýva, že rok 2020 by mal byť o niečo reprezentatívnejší ako rok 2019, práve kvôli vyššiemu počtu platných dát na väčšine územia.



Obr.1: Počet platných dát v roku 2019 (naľavo) a 2020 (napravo) a rozdiel v počte platných dát (dole).

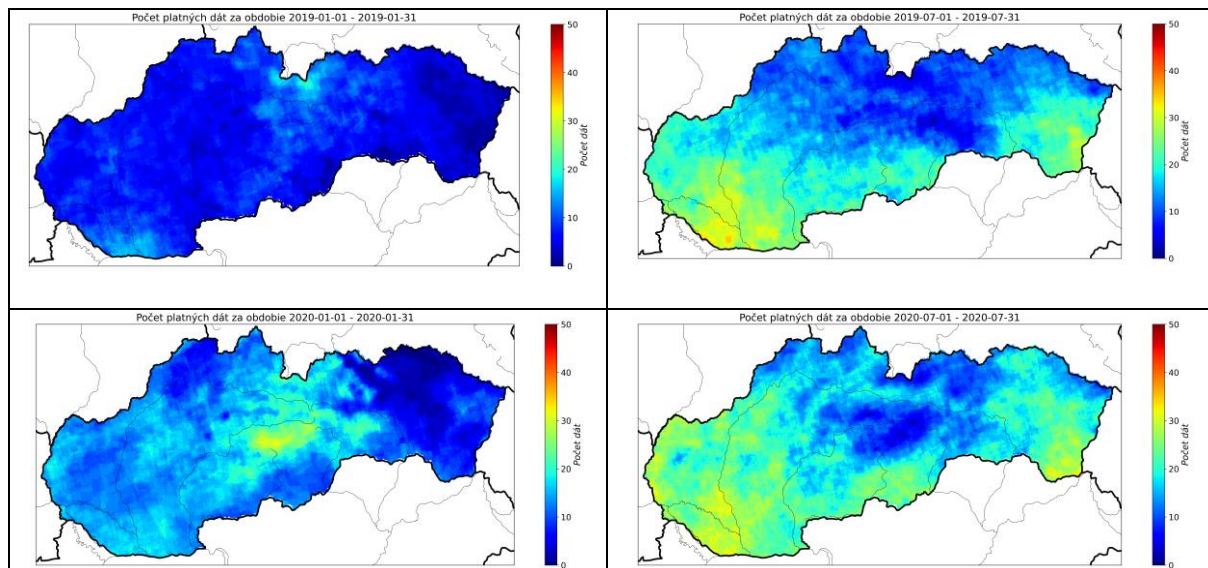
Okrem počtu platných dát nás zaujíma tiež čas počas dňa, kedy bola družicová snímka vykonaná. V tomto prípade sú dáta k dispozícii každý deň 2 krát doobeda (niekedy len 1 krát), čo môžeme považovať za výrazný limitujúci faktor, keďže dáta z pozemnej monitorovacej siete máme k dispozícii kontinuálne počas dňa. Z histogramu platných údajov nad RB stanicami (Obr. 1.1) vidno počet dát v jednotlivých časových intervaloch.



Obr.1.1: Histogram počtu platných dát v jednotlivých časových intervaloch nad RB stanicami.

Ďalším pozorovaním je počet kvalitných dát počas roka. Ukázalo sa, že najviac dát je na jar a v lete, avšak vtedy je priemerná hodnota troposférického stĺpca najnižšia. Naopak najmenej dát je v zime, lenže vtedy je priemerná hodnota troposférického stĺpca najvyššia. Pokiaľ porovnáme počet dát pre jednotlivé mesiace v roku, je pomerne dobre vidno rozdiel medzi januárom a júlom. V januári je viac dát v oblasti hôr, ktoré sa nachádzajú často nad inverziou a nízkou oblačnosťou. V júli je opačná

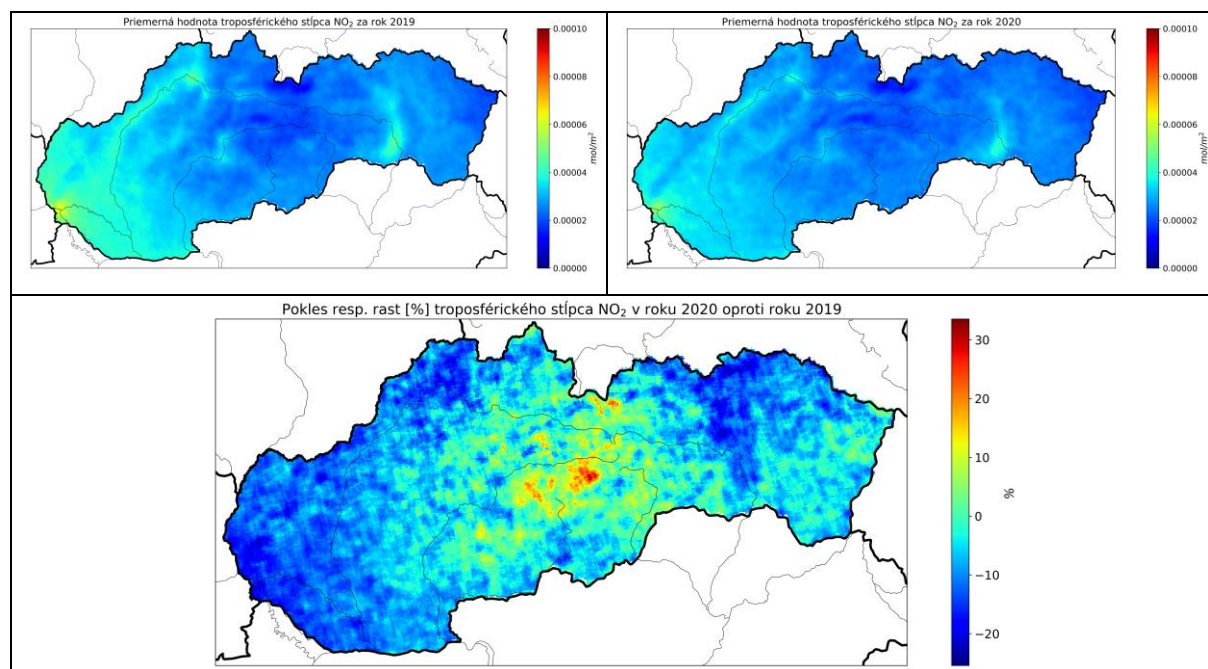
situácia a menej dát je v hornatých oblastiach stredného Slovenska kvôli zväčšenej oblačnosti. Vyššie spomenutý rozdiel medzi januárom a júlom pozorujeme pre oba roky 2019 aj 2020 (Obr.2).



Obr.2: Počet platných dát v januári (naľavo) a v júli (napravo) pre rok 2019 (hore) a rok 2020 (dole)

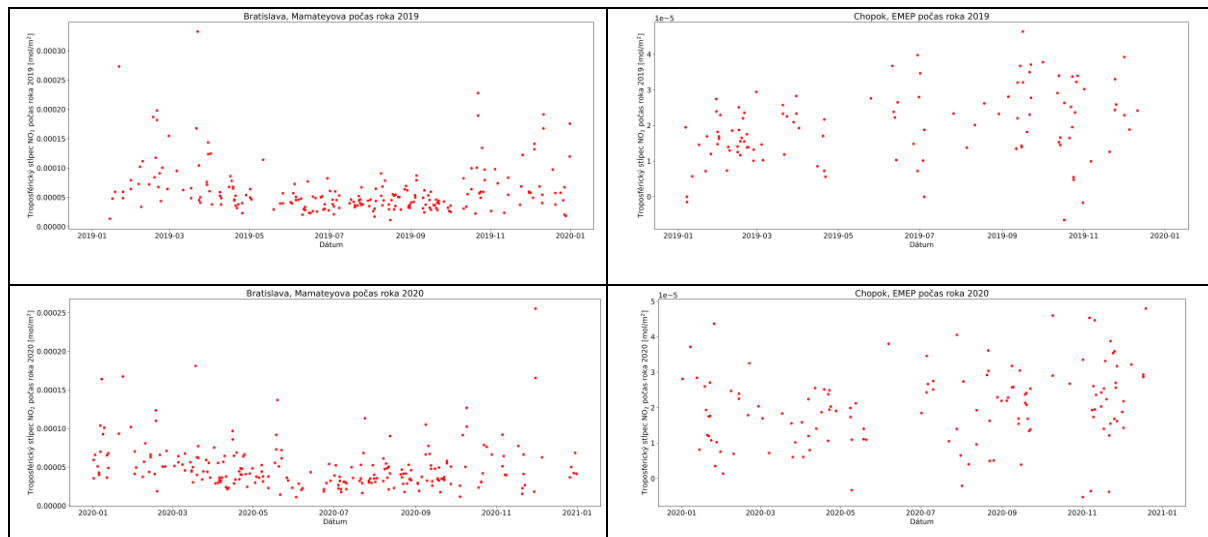
Priestorová a časová analýza rozloženia priemernej hodnoty troposférického stĺpca pre oblasť SR

Na Obr.3 je zobrazená priemerná hodnota troposférického stĺpca NO_2 na území SR za roky 2019 a 2020. Ďalej je zobrazená percentuálna zmena v roku 2020 oproti roku 2019 podľa vzťahu $((2020-2019)/2019)*100\%$. V roku 2020 došlo k zníženiu priemernej hodnoty troposférického stĺpca najmä na západe, severozápade a východe Slovenska miestami aj o 20%. Toto zníženie sa dá pripísať zníženej mobilite obyvateľstva v období pandémie. Naopak vzrast hodnôt je na hornatom území stredného Slovenska. Tento vzrast je tu zrejme spôsobený náhodnou fluktuáciou dát, keďže v tejto oblasti je málo platných dát a hodnoty troposférického stĺpca sú tu veľmi nízke.



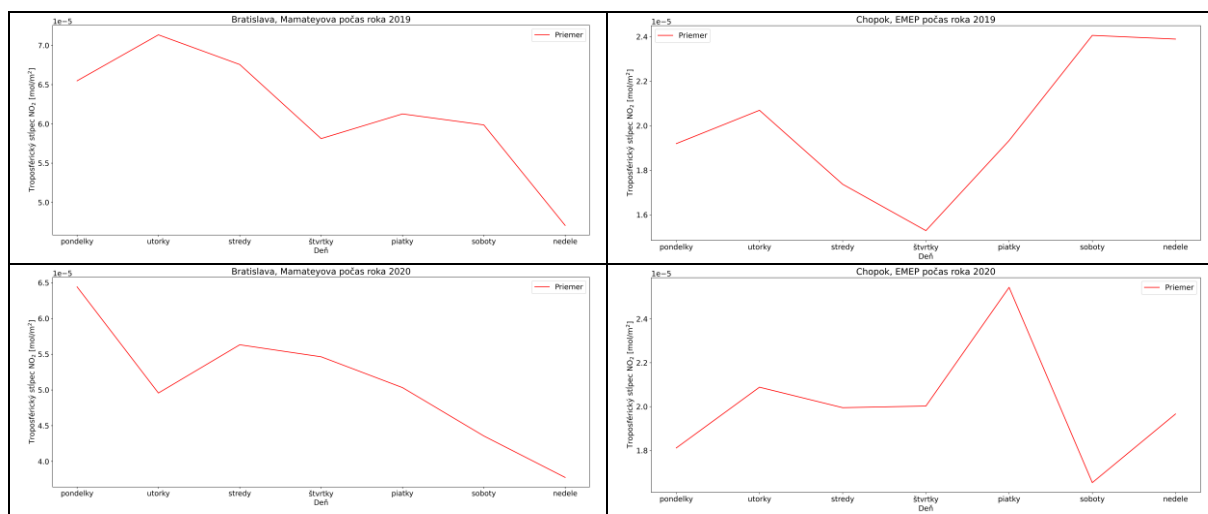
Obr.3: Priemerná hodnota troposférického stĺpca za rok 2019 (naľavo) a 2020 (napravo) a percentuálny rast alebo pokles v roku 2020 oproti roku 2019 (dole)

V mestách a v oblastiach so zvýšenou hustotou obyvateľstva má troposférický stĺpec typický ročný chod s minimom v lete a maximom v zimnom období. Na horách ročný chod nebol zaznamenaný, teda hodnoty počas roka sa menia iba minimálne alebo vôbec. Ako príklad uvádzame hodnoty troposférického stĺpca počas rokov 2019 a 2020 nad stanicami Bratislava, Mamateyova a Chopok, EMEP (Obr. 3).

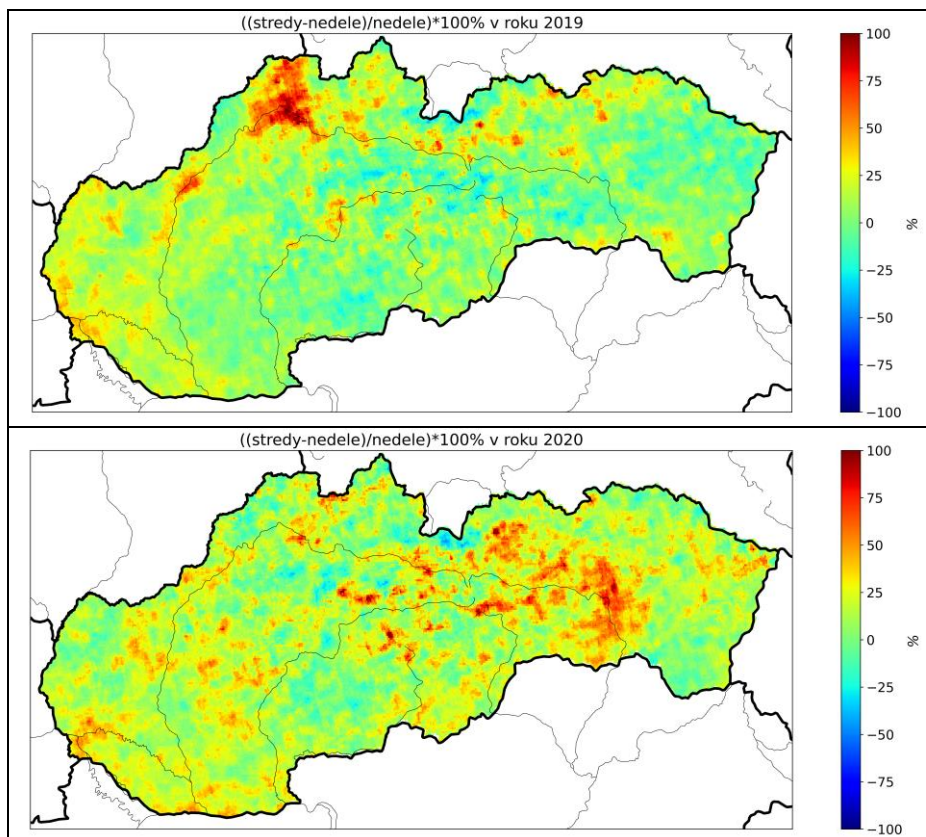


Obr.3: Bratislava (naľavo) a Chopok (napravo) pre rok 2019 (hore) a rok 2020 (dole)

Pokiaľ ide o porovnanie jednotlivých dní v týždni, pre mestské oblasti pozorujeme pokles cez víkend a nárast v pracovné dni, pričom tento rozdiel je na úrovni cca. 1 mol/m^2 . Nad stanicou Chopok EMEP, kvôli nízkemu počtu dát a tiež nízkym hodnotám troposférického stĺpca NO_2 , zreteľný trend nevidieť (Obr.4). Pokiaľ zoberieme priemernú hodnotu za všetky stredu a všetky nedele (streda reprezentuje stred pracovného týždňa a zvýšenú mobilitu a nedeľa zníženú mobilitu), môžeme podľa vzťahu: $((\text{streda-nedeľa})/\text{nedeľa}) * 100 \%$ získať percentuálny nárast alebo pokles hodnôt za stredu oproti hodnotám v nedeľu za daný rok (Obr.5). Poznamenajme, že počet údajov je najmä v horských oblastiach príliš nízky.



Obr.4: Bratislava (naľavo) a Chopok (napravo) pre rok 2019 (hore) a rok 2020 (dole) pre jednotlivé dni v týždni



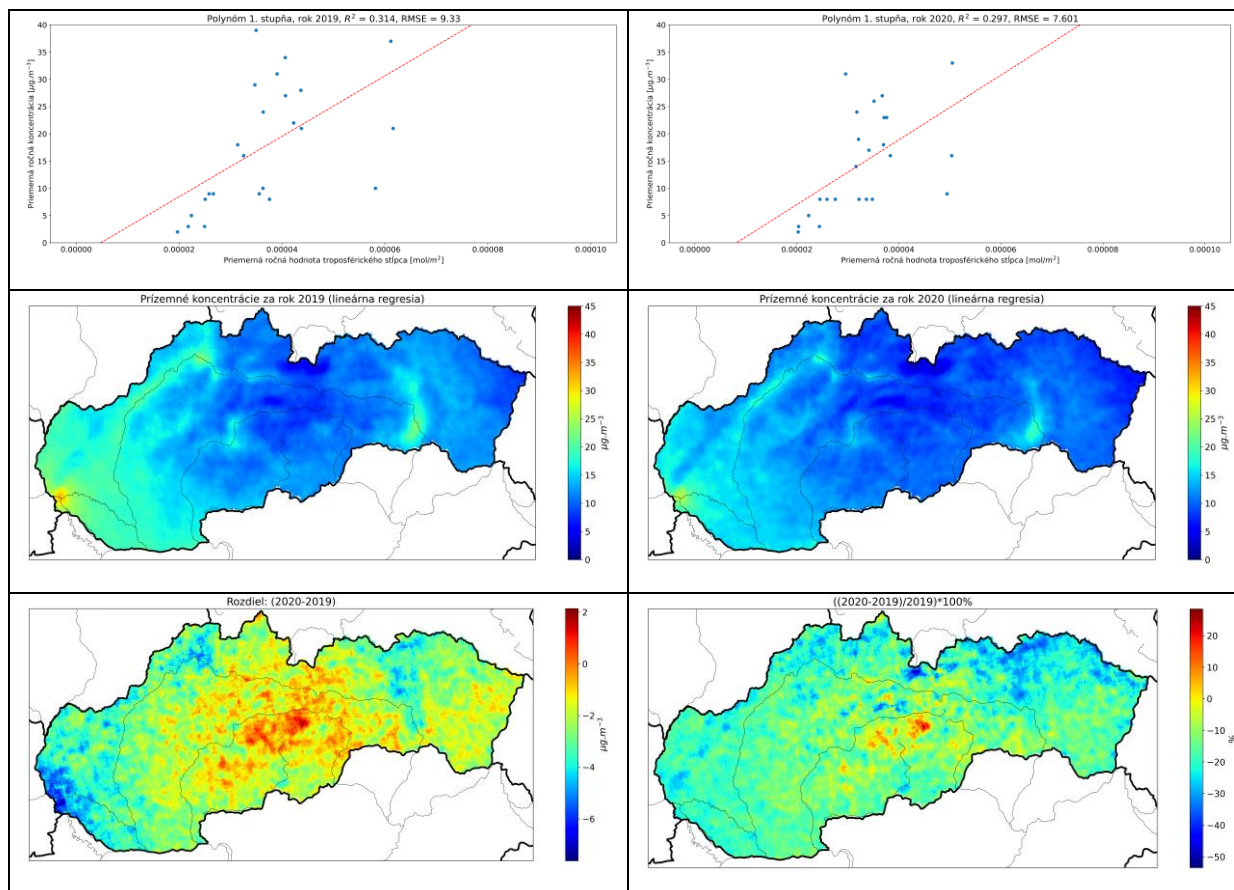
Obr.5: Porovnanie hodnôt za stredy a nedele v rokoch 2019 a 2020

Vzťah družicových meraní s pozemnými meraniami koncentrácií NO₂

Ďalšou otázkou je vzťah nameraného troposférického stĺpca NO₂ pomocou družice Sentinel-5P s nameranými koncentraciami NO₂ na pozemných staniciach NMSKO. Pokiaľ za nezávislú premennú x zvolíme merania z družice (priemerné ročné hodnoty troposférického stĺpca NO₂) a závislú premennú y staničné merania NMSKO (priemerné ročné koncentrácie), môžeme takto vytvoriť regresnú predpoveď prízemných koncentrácií pre celé územie SR. Takýto lineárny regresný model má tvar $y = ax + c + \varepsilon$, kde a , c sú vypočítané parametre a ε je chyba modelu oproti reálne nameranej hodnote pomocou pozemných meraní. Týmto spôsobom odhadneme priemernú ročnú prízemnú koncentráciu na celom území SR len na základe predpokladanej lineárnej závislosti medzi družicovými a pozemnými meraniami (Obr.6). Parametre lineárnej regresie sú uvedené v Tab.1.

Tab.1: Parametre regresnej priamky (polynómu 1. stupňa)

Parameter	Hodnota parametra (rok 2019)	Hodnota parametra (rok 2020)
c	-2,62	-4,84
a	554459,74	593882,96
r ²	0,31	0,29
RMSE	9,33 µg/m ³	7,601 µg/m ³

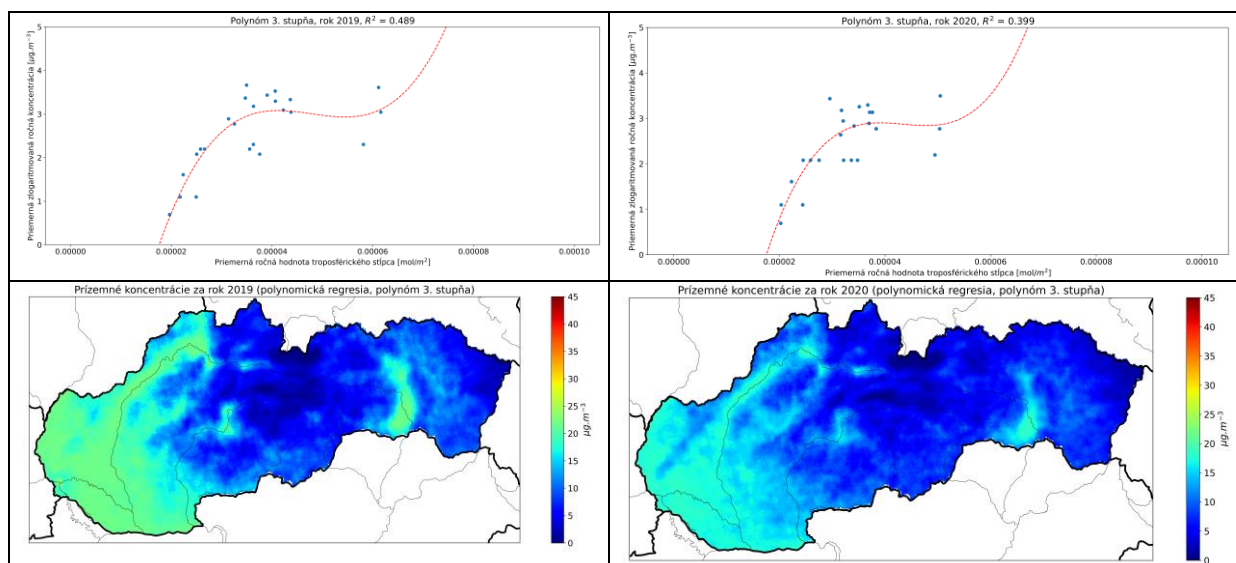


Obr.6: Vrchné obrázky: Rozptylové grafy medzi priemernými ročnými hodnotami troposférického stĺpca NO_2 a koncentraciami NO_2 nameranými na staniách NMSKO. Červená čiara predstavuje regresnú priamku. Obrázky v strede: Hodnoty prízemných koncentrácií získané pomocou regresie z priemerných ročných hodnôt troposférického stĺpca NO_2 pre celé územie SR. Obrázky dole: Rozdiel prízemných koncentrácií.

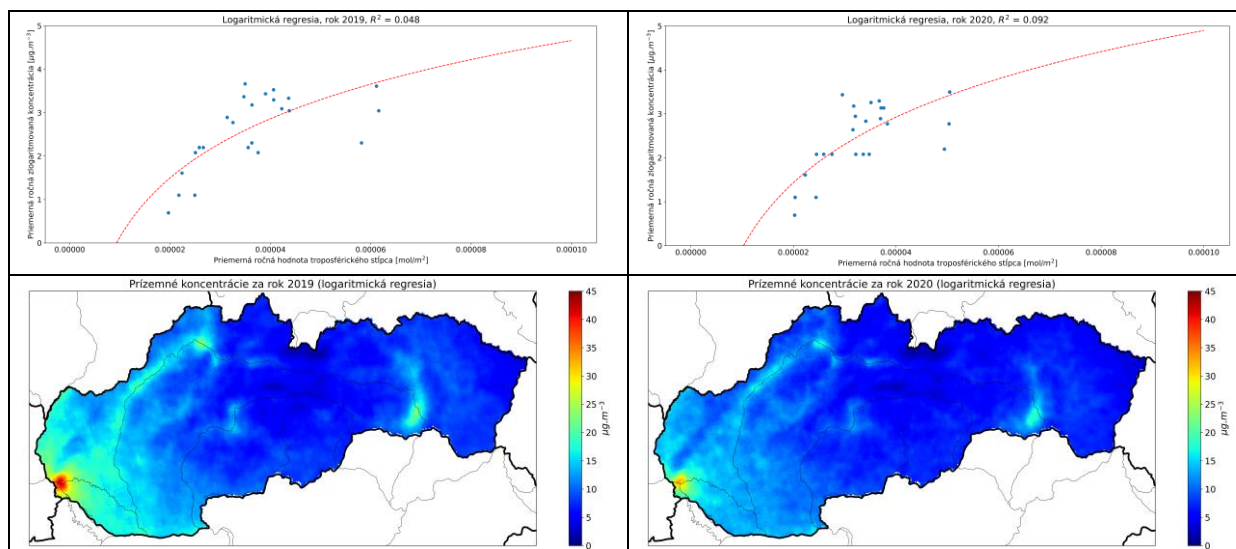
Stupeň príčinnej závislosti premennej y od premennej x vyjadruje koeficient determinácie, definovaný ako druhá mocnina koeficientu korelácie r . Vo výberovom súbore ho označujeme ako r^2 . Interpretácia koeficientu determinácie vychádza z analýzy variability (rozptylu) závisle premennej y , ktorú by mala do značnej miery vysvetliť variabilita nezávisle premennej x za predpokladu, že od nej lineárne závisí veľkosť hodnôt y [6]. V našom prípade je $r^2 = 0,31$, čo znamená, že iba 31% variability premennej y sa dá vysvetliť lineárnym vzťahom s premennou x (regresnou priamkou). To môže byť spôsobené tým, že družicové merania majú pomerne hrubé rozlíšenie, poskytujú merania len 1 až 2 krát denne a merajú NO_2 v celom stĺpci troposféry, ale aj tým, že lineárny vzťah medzi družicovými a pozemnými meraniami je nevhodný.

Za účelom vystihnúť lepšej závislosti medzi družicovými a pozemnými meraniami sme vyskúšali viacero polynomiálnych závislostí a logaritmickú závislosť v tvare $y = a \cdot \ln(x) + c + \varepsilon$. Zvyšovanie stupňa polynómu na väčší ako 3. stupeň sa ukázalo ako nevhodné kvôli zbytočnému prekmitávaniu polynómu medzi bodmi merania (overfitting). Aby regresná predpoveď negenerovala záporné prízemné koncentrácie, hodnoty závisle premennej y sme najprv zlogaritmovali, následne vykonali regresiu a závislú premennú sme spätne exponenciálne transformovali. Výsledné mapy prízemných koncentrácií sú na Obr. 7 a 8. Ako vidno z Obr.7 a 8, hoci je štatisticky polynomiálny fit 3. stupňa lepší (vyššie r^2) ako lineárny, jeho regresná krivka je pomerne nerealistická, čo sa v konečnom dôsledku prejavilo v menšej variabilite vypočítaných prízemných koncentrácií, pričom najviac znečistené oblasti NO_2 tu nevidno – napr. okolie Bratislavy. RMSE pre rok 2019 vyšlo $8,049 \mu\text{g}/\text{m}^3$ a pre rok 2020 vyšlo $7,02 \mu\text{g}/\text{m}^3$. V prípade logaritmického fitu sa nám lepšie zvýraznili hot-spoty, ale

výsledná validácia je horšia ako pre lineárny fit (nižšie r^2) a $RMSE=10,98 \mu\text{g}/\text{m}^3$ pre rok 2019 a $RMSE=8,64 \mu\text{g}/\text{m}^3$ pre rok 2020.



Obr.7: Vrchné obrázky: Rozptylové grafy medzi priemernými ročnými hodnotami troposférického stĺpca NO_2 a zlogaritmovanými koncentraciami NO_2 nameranými na stanicach NMSKO. Červená čiara predstavuje polynóm 3. stupňa. Spodné obrázky: Hodnoty prízemných koncentrácií získané pomocou polynomickej regresie z priemerných ročných hodnôt troposférického stĺpca NO_2 pre celé územie SR.



Obr.8: Vrchné obrázky: Rozptylové grafy medzi priemernými ročnými hodnotami troposférického stĺpca NO_2 a zlogaritmovanými koncentraciami NO_2 nameranými na stanicach NMSKO. Červená čiara predstavuje logaritmickú funkciu. Spodné obrázky: Hodnoty prízemných koncentrácií získané pomocou logaritmickej regresie z priemerných ročných hodnôt troposférického stĺpca NO_2 pre celé územie SR.

Záver

Družica Sentinel-5P poskytuje dáta, ktoré majú vysoký potenciál prispieť ku skvalitneniu monitorovania kvality ovzdušia. Najväčšou výhodou družicových dát je, že na rozdiel od chemicko-transportných modelov neobsahujú veľké množstvo vstupných parametrov, ktoré sú zaťažené neistotami. Nevýhodou je nízky počet dát (cca. 2 snímky za deň), ktoré sú navyše výrazne ovplyvnené oblačnosťou, a tiež to, že družica nesníma prízemné koncentrácie, ale obsah NO₂ v celom stĺpci atmosféry, čo prirodzene vedie ku hľadaniu vzťahu medzi týmito dvoma veličinami. Na základe skúmania viacerých vzťahov medzi družicovými a pozemnými meraniami boli tiež pre dané roky vytvorené mapy prízemných koncentrácií NO₂. Predpokladáme, že lineárna závislosť vystihuje vzťah medzi družicovými a pozemnými meraniami najlepšie ($r = 0,56$ pre rok 2019 a $r = 0,54$ za rok 2020), aj keď pre konečné posúdenie je potrebná podrobnejšia validácia. Bolo zistené zníženie koncentrácií v roku 2020 oproti 2019 na väčšine územia Slovenska, čo môže byť spôsobené zmenou správania sa spoločnosti súvisiacou s pandémiou COVID-19.

3 Literatúra

- [1]<https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-5p> (posledný prístup 21.10.2021)
- [2]<https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Nitrogen-Dioxide-Level-2-Product-Readme-File> (posledný prístup 21.10.2021)
- [3] Schneider, Philipp & Hamer, Paul & Kylling, Arve & Shetty, Shobitha & Stebel, Kerstin. (2021). Spatiotemporal Patterns in Data Availability of the Sentinel-5P NO₂ Product over Urban Areas in Norway. *Remote Sensing*. 13. 2095. 10.3390/rs13112095.
- [4] Kaplan, Gordana & Yigit Avdan, Zehra. (2020). Space-Borne Air Pollution Observation from Sentinel-5p TROPOMI: Relationship Between Pollutants, Geographical and Demographic Data. 10.26833/ijeg.644089.
- [5] Oxoli, Daniele & Cedeno, Rodrigo & Brovelli, Maria. (2020). ASSESSMENT OF SENTINEL-5P PERFORMANCE FOR GROUND-LEVEL AIR QUALITY MONITORING: PREPARATORY EXPERIMENTS OVER THE COVID-19 LOCKDOWN PERIOD. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLIV-3/W1-2020. 10.5194/isprs-archives-XLIV-3-W1-2020-111-2020.
- [6] Oršanský, P., & Ftorek, B. (2017). *Štatistické a numerické metódy: Učebný text* (Vydanie prvé.). Žilina: Edis - vydavateľské centrum ŽU.
- [7] Zhao, X., Griffin, D., Fioletov, V., McLinden, C., Cede, A., Tiefengraber, M., Müller, M., Bognar, K., Strong, K., Boersma, F., Eskes, H., Davies, J., Ogyu, A., and Lee, S. C.: Assessment of the quality of TROPOMI high-spatial-resolution NO₂ data products in the Greater Toronto Area, *Atmos. Meas. Tech.*, 13, 2131–2159, <https://doi.org/10.5194/amt-13-2131-2020>, 2020.

English Abstract:

In this paper, data availability and the tropospheric NO₂ column in the territory of Slovakia from the Sentinel-5P satellite was analysed. The data were from years 2019 a 2020. The quality of the data depends on clouds, surface albedo, the presence of ice and snow. The data quality is characterized by the qa value, which can take a continuous value from 0 to 1. The qa value represents the limit value of the individual quality of each pixel. In the case of the tropospheric NO₂ column the recommended value is $qa > 0,75$. The data are available every day 2 times in the morning (sometimes only once), which can be considered a significant limiting factor, as the data from the terrestrial monitoring network are available continuously during the day. In 2019, the total number of satellite images over the Slovak republic was 576 and in 2020 it was 569 images. The total number of valid data in the specific location during 2019 were in the range from 103 to 272 depending on the location. In 2020, the number of valid data ranged from 101 to 264. The minimum number of valid data was in the

mountainous areas of central Slovakia and maximum number in the lowlands, where it is the least cloudy during the year.

Another observation is the number of quality data during the year. It turned out that most data are in spring and summer when the average value of the tropospheric column is the lowest. On the contrary, minimum of data is available in winter, when the average value of the tropospheric column is the highest. For urban areas, we observe a decrease over the weekend and an increase on weekdays.

In cities and areas with increased population density, the tropospheric column has a typical annual run with a minimum in summer and a maximum in winter. In the mountain areas, the annual course was not recorded.

The relationship between ground based observations and observations from the satellite was analysed as well. There is a positive correlation $r = 0.56$ for 2019 and $r = 0.54$ for 2020 between ground based observations and observations from satellite. The annual mean of ground concentrations was calculated for the territory of Slovakia for years 2019 and 2020 by linear regression model and was compared between the years. It appears that the logarithmic and polynomial regression models were not suitable for estimation of ground NO₂ concentrations from the satellite NO₂ tropospheric column measurements.