



Vlaanderen
is wetenschap



12_076_12
WL rapporten

Data import sedimentconcentraties en rapportage

Deelrapport 2 – correctie YSI data

DEPARTEMENT
MOBILITEIT &
OPENBARE
WERKEN

waterbouwkundiglaboratorium.be

Data import sedimentconcentraties en rapportage

Deelrapport 2 – correctie YSI data

De Decker, K.; Leyssen, G.; Viaene, P.; Vanlierde, E.; Deschamps, M.; Verwaest, T.; Mostaert, F.

Juridische kennisgeving

Het Waterbouwkundig Laboratorium is van mening dat de informatie en standpunten in dit rapport onderbouwd worden door de op het moment van schrijven beschikbare gegevens en kennis.

De standpunten in deze publicatie zijn deze van het Waterbouwkundig Laboratorium en geven niet noodzakelijk de mening weer van de Vlaamse overheid of één van haar instellingen.

Het Waterbouwkundig Laboratorium noch iedere persoon of bedrijf optredend namens het Waterbouwkundig Laboratorium is aansprakelijk voor het gebruik dat gemaakt wordt van de informatie uit dit rapport of voor verlies of schade die eruit voortvloeit.

Copyright en wijze van citeren

© Vlaamse overheid, Departement Mobiliteit en Openbare Werken, Waterbouwkundig Laboratorium 2016
D/2016/3241/313

Deze publicatie dient als volgt geciteerd te worden:

De Decker, K.; Leyssen, G.; Viaene, P.; Vanlierde, E.; Deschamps, M.; Verwaest, T.; Mostaert, F. (2016). Data import sedimentconcentraties en rapportage: Deelrapport 2 – correctie YSI data. Versie 3.0. WL Rapporten, 12_076_12. Waterbouwkundig Laboratorium: Antwerpen.

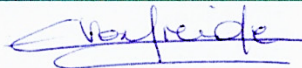
Overname uit en verwijzingen naar deze publicatie worden aangemoedigd, mits correcte bronvermelding.

Documentidentificatie

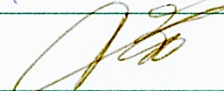
Opdrachtgever:	Waterbouwkundig Laboratorium	Ref.:	WL2016R12_076_12
Keywords (3-5):	ISCO, LIMS, WISKI		
Tekst (p.):	17	Bijlagen (p.):	/
Vertrouwelijk:	☒ Nee	☒ Online beschikbaar	

Auteur(s):	De Decker, K.; Leyssen, G.
------------	----------------------------

Controle

	Naam	Handtekening
Revisor(en):	Viaene, P.	
Projectleider:	Vanlierde, E.	

Goedkeuring

Verantwoordelijke (Steunpunt) HIC:	Deschamps, M.	
Coördinator onderzoeksgroep:	Verwaest, T.	
Afdelingshoofd:	Mostaert, F.	



Abstract

Dit rapport beschrijft hoe problemen met YSI data (opgemerkt in het voorjaar van 2016, waarbij de data in de databank niet overeenkomt met de data aanwezig in de YSI interne logging) kunnen rechtgezet worden a.d.h.v. scripts opgesteld hiervoor.

Inhoudstafel

Abstract	III
Inhoudstafel.....	V
Lijst van de tabellen.....	VII
1 Probleemschetsing	1
1.1 Inleiding	1
1.2 Overzicht meetposten	1
2 Data	2
2.1 Data vanuit telemetrie	2
2.2 Data in interne logbestanden YSI	2
3 Verkennende analyse	3
4 Methodologie correcties	4
4.1 Werking en ontwikkeling scripts	4
4.1.1 Omkaderingsscripts	4
4.1.2 Aanstuurscript:	4
4.1.3 Hoofdscrip:	5
4.2 Methodologie YSI analyse	6
4.2.1 Fase 1: zoeken van eerste ‘match’:	6
4.2.2 Fase 2: zoeken van overeenkomsten voor de rest van de reeks.....	6
4.3 Toekomstige correcties: 2016	7
4.3.1 Export van WISKI-O reeksen	7
4.3.2 YSI – interne log bestanden	7
4.3.3 Runnen master_YSI_frombatch.py.....	8
5 YSI-data correctie	9
5.1 Prioritaire posten.....	10
5.1.1 Melle	10
5.1.2 Dendermonde.....	11
5.1.3 Duffel	11
5.1.4 Eppegem	12
5.2 Niet-prioritaire posten.....	13
5.2.1 Kruibeke.....	13
5.2.2 Aarschot.....	13

5.2.3	Destelbergen	14
5.2.4	Gavere	14
5.2.5	Grobbendonk.....	15
5.2.6	Itegem.....	15
5.2.7	Lippenbroek_in.....	15
5.2.8	Lippenbroek_uit	15
5.2.9	Wondelgem	16
5.2.10	Zulte.....	16
5.2.11	Lummen.....	17
5.2.12	Halen.....	17

Lijst van de tabellen

Tabel 1: Overzicht meetposten	1
Tabel 2: Locaties van analyseplots.	9
Tabel 3: Afwijkingen tussen interne log en WISKI reeks.	10
Tabel 4: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	10
Tabel 5: Afwijkingen tussen interne log en WISKI reeks.	11
Tabel 6: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	11
Tabel 7: Afwijkingen tussen interne log en WISKI reeks.	11
Tabel 8: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	12
Tabel 9: Afwijkingen tussen interne log en WISKI reeks.	12
Tabel 10: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	12
Tabel 11: Afwijkingen tussen interne log en WISKI reeks.	13
Tabel 12: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	13
Tabel 13: Afwijkingen tussen interne log en WISKI reeks.	13
Tabel 14: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	14
Tabel 15: Afwijkingen tussen interne log en WISKI reeks.	14
Tabel 16: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	14
Tabel 17: Afwijkingen tussen interne log en WISKI reeks.	14
Tabel 18: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	15
Tabel 19: Afwijkingen tussen interne log en WISKI reeks.	15
Tabel 20: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	15
Tabel 21: Afwijkingen tussen interne log en WISKI reeks.	16
Tabel 22: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	16
Tabel 23: Afwijkingen tussen interne log en WISKI reeks.	16
Tabel 24: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.	17

1 Probleemschetsing

1.1 Inleiding

Na een verkennende vergelijking van Andy Henderick ([P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\analyse_andy\Controle Tijdsverschuiving.xlsx](#)) is aan het licht gekomen dat de data doorgestuurd via telemetrie vanaf de YSI meetinstrumenten op bepaalde momenten sterk afwijkt van de waarden die in de interne log bestanden van de YSI worden weggeschreven. Dit probleem is in dit document verder onderzocht en er is een correctie van de meetdata uitgevoerd in de WISKI-database. Voor deze correctie is de ruwe data uit de interne logbestanden van de YSI gebruikt. Aangezien geweten is dat de interne klok van de YSI een afwijking heeft, wordt de data nog verschoven in de tijd aan de hand van een correlatie-analyse met de beschikbare meetdata in WISKI (telemetrie).

De scripts die geschreven werden voor de correctie van YSI werden geschreven in Python versie 2.7. De Python-distributie die daarvoor gebruikt werd is Anaconda (installatie van Python met inbegrip van een aantal wetenschappelijke packages die nodig zijn voor het uitvoeren van de scripts). De IDE die gebruikt werd om de scripts te ontwikkelen (grafische user interface voor python, waar ontwikkeling en debugging van de scripts mogelijk is) is PyCharm.

1.2 Overzicht meetposten

Hier is een overzicht gegeven van de 16 meetposten waarvoor de controle en correctie noodzakelijk is.

Tabel 1: Overzicht meetposten

Locatie	Code	opmerking
Aarschot	Dem02a	
Dendermonde	Den02a	prioritair
Destelbergen	Zes57n	
Duffel Sluis	Bnt03a	prioritair
Eppegem	Zen03a	prioritair
Gavere	Bos02a	
Halen	Dem12a	
Lippenbroek UIT	Zes40a	
Lippenbroek IN	Zes40d	
Lummen	Dem15a	
Melle	Zes57a	prioritair
Wondelgem	Rvg02a	
Zulte	Lei05a	
Grobbendonk	Knt03a	
Itegem Hullebrug	Gnt03a	
Kruikeke	Zes24a	

2 Data

2.1 Data vanuit telemetrie

De data binnengekomen via telemetrie is in de WISKI6 databank opgeslagen. Er is een WISKI export gemaakt voor alle variabelen gemeten door de YSI meetinstrumenten van zowel de o-reeks als de p-reeks.

- O-reeksen: P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\WISKI\WL_2015_26\Uitvoering\YSI\WISKI_preeksen\
- P-reeksen: P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\WISKI_preeksen\

2.2 Data in interne logbestanden YSI

De YSI-bestanden werden van het binaire *.dat-formaat naar ASCII formaat omgezet, gebruikmakend van het programma EcoWatch Lite. De *.txt-bestanden zijn beschikbaar op P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\YSI_bestand\. De gemeten en geëxporteerde parameters zijn de volgende:

Parameter	Eenheid
Date	d/m/yy d/m/yyyy m/d/yy m/d/yyyy
Time	HH:MM:SS
Temp	°C
SpCond	µS/cm
Cond	µS/cm
Press	psia
pH	pH
pHmV	pH mV
Orp	mV
Turbid+	NTU
ODOsat	%
ODO	mg/l
Battery	volts

3 Verkennde analyse

Voor elke YSI-post is een vergelijking gemaakt tussen de YSI-data in de interne logger van de YSI en de data in de O-reeks van WISKI. De YSI-data is beschikbaar in verschillende bestanden. Tussen de data in deze bestanden en de data in de WISKI-tijdreeks zit er echter een tijdsverschuiving, aangezien de interne klok van de YSI-meettoestellen een afwijking bezit. Deze tijdsverschuiving verschilt van bestand tot bestand en gaat van 5 minuten tot maximaal 2 uur. Doormiddel van een correlatie optimalisatie van een discreet aantal tijdsverschuivingen is een optimale verschuiving bepaald per YSI-bestand. Hoewel geweten is dat de tijdsverschuiving binnen een YSI-bestand kan variëren wordt deze correctie als voldoende beschouwd voor de verkennende analyse. In de werkelijke correctie worden deze extra moeilijkheden wel mee in rekening gebracht. Voor elke post is per jaar een interactieve html figuur gemaakt met 3 assen. De bovenste figuur geeft de twee tijdreeksen weer (YSI en WISKI-O), de middelste figuur geeft het verschil in turbiditeit groter dan 1 weer in log schaal en de onderste het nummer van het YSI meetinstrument. De figuren zijn terug te vinden onder: P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\+ **naam van de post** + \diff_plots. Voor deze analyse werd het script \bin\paramplot.py gebruikt met de parameter 'diff', aan te roepen met: **cmd** *python paramplot.py Melle Turbid diff 2015*.

Op basis van deze verkennende analyse wordt beslist welke meetposten gecorrigeerd moeten worden. De prioritaire posten (Tabel 1) dienen zeker gecorrigeerd te worden.

4 Methodologie correcties

4.1 Werking en ontwikkeling scripts

Locatie code: [P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\bin\](#)

De code bestaat uit drie onderdelen: de werkelijke code, bestaande uit analyse van het probleem en correctie van enerzijds de WISKI originele (O) reeks, anderzijds de WISKI productie (P) reeks; een aanstuurscript die de code kan aansturen en parameters aanvaardt; enkele omkaderingsfuncties die voor de pre- en postprocessing van de input data zorgen. Hieronder wordt telkens een voorbeeld gegeven van hoe het script te runnen in command line mode (**CMD**).

4.1.1 Omkaderingsscripts

Clip_wiski.py: dit script kan gebruikt worden om de geëxporteerde WISKI-data per jaar op te delen en in aparte mappen te plaatsen. Dit versnelt de import van WISKI-data tijdens de analyse, correctie en visualisatie.

CMD (1 jaar): `python clip_wiski.py Melle 2015 o` of `python clip_wiski.py Melle 2015 p`

CMD (alle jaren): `for %s in (2015,2014,2013,2012,2011) do start python clip_wiski.py Melle %s o`

Paramplot.py: dit script bevat verschillende functies om de geëxporteerde zrx bestanden na analyse te verwerken tot een dynamische html plot. Het type plot wordt gestuurd door een parameter in de code.

- 'a' (all): visualiseert originele WISKI P, WISKI O, YSI, gecorrigeerde WISKI O en gecorrigeerde WISKI P.

CMD: `python paramplot.py Melle Turbid a 2015`

- 'diff' (difference): snelle analyse van de verschillen tussen WISKI O en YSI op basis van één correlatie per YSI-bestand. Geeft WISKI O, YSI, verschil en YSI-toestelnummer weer.

CMD: `python paramplot.py Melle Turbid diff 2015`

4.1.2 Aanstuurscript:

master_YSI_frombatch.py: dit script bevat de beslissingslogica om de verschillende delen van het hoofdsript aan te roepen: keuze tussen correctie van WISKI O en/of P reeks per jaar, en voor 1 parameter of alle parameters. Indien de analyse reeds is gebeurd voor turbiditeit wordt meteen aan de correctie begonnen.

`python master_YSI_frombatch.py [post] [param] [o/p/all] [jaar]`

CMD (1) `python master_YSI_frombatch.py Melle Turbid o 2015`

CMD (2): `python master_YSI_frombatch.py Melle all all 2015`

CMD (3): `for %s in (2015,2014,2013,2012,2011) do start python master_YSI_frombatch.py Melle all all %s`

Opmerking: Indien men slechts de eerste fase wil uitvoeren kan men in dit script (zie lijn 57) enkele lijnen uitcommentariëren.

4.1.3 Hoofdsript:

YSI analyse_obj.py: dit script bevat de methodologie van de YSI analyse en correctie van de WISKI-O en –P reeksen. Het script bevat drie functies (parameters worden tussen haakjes gegeven):

- Master YSI analysis(*workdir, YSI_dir, WISKI_dir, param="Turbid", time="2015"*): dit script bevat de analyse van de interne log (YSI) bestanden. Input bestanden zijn de WISKI-O reeks van de post/parameter/jaar combinatie en telkens 1 YSI *.txt bestand. Doel is zo goed als mogelijk een overeenkomst te vinden tussen de YSI reeks en WISKI-O reeks. Dit wordt bemoeilijkt door enerzijds een tijdfout op de YSI-bestanden (interne klok YSI-meettoestel loopt achter), en anderzijds een foute overdracht van gegevens tussen YSI-meettoestel en WISKI-databank. De methode is daarom gebaseerd op een flexibele correlatie-analyse. Output is een tekst bestand waarin voor elk YSI tijdstip het overeenkomende WISKI tijdstip is weggeschreven, met telkens alle YSI en WISKI parameter waarden van de periode. De methodologie is uitgewerkt in §4.2.
- Correction wiski_original(*workdir, parameter, time*): Na de analysestap wordt het resulterende *.txt bestand aangewend om een nieuwe WISKI-O zrx aan te maken. Voor elke parameter wordt een nieuw dataframe opgebouwd bestaande uit de WISKI-O tijdsreeks enerzijds en de gemeten YSI-waarden anderzijds (YSI tijdstippen bevatten een tijdverschuiving, WISKI waarden een logfout). Dit nieuw dataframe wordt dan weggeschreven naar een nieuw ZRX-bestand, gebruikmakend van de ZRX-headers van het originele WISKI-O ZRX-bestand.
- Correction wiski_production(*workdir, WISKI_p_dir, WISKI_o_dir, parameter, time*): dit script corrigeert de WISKI-P reeks op basis van de output van master_YSI_analysis. Allereerst wordt op basis van de standaarddeviatie van de WISKI-P reeks een tolerantie bepaald (1% stdev). Daarna wordt de WISKI-P reeks chronologisch doorlopen, en wordt telkens de bijhorende commentaar gelezen, om te zien of een tijdsverschuiving werd toegepast. Indien het script zo'n tijdverschuiving herkent (syntax '**10 min** opgeschoven naar **voor**'), wordt de omgekeerde verschuiving toegepast op de geselecteerde timestamp. Er wordt dan in de analyse-output gezocht naar een overeenkomende WISKI-O timestamp, waardoor meteen de werkelijk gemeten YSI waarde voor dit moment gekend is. Vervolgens wordt volgens onderstaande regels beslist wat met de nieuwe waarde gedaan moet worden:
 - Indien de YSI-waarde afwijkt van de WISKI-O waarde met meer dan de vooraf bepaalde tolerantie, wordt de YSI waarde overgenomen op het WISKI-P tijdstip.
 - Indien de YSI-waarde niet meer dan de tolerantie afwijkt van de WISKI-O waarde blijft de WISKI-P-waarde behouden.
 - Indien de WISKI-P-waarde op missing value werd gezet (en dus de WISKI-O-waarde geen missing value was), wordt de YSI-waarde niet overgenomen.
 - Indien én de WISKI-P-waarde én de WISKI-O-waarde missing values zijn (en dus de data niet correct naar de WISKI-databank werd doorgezonden), wordt de YSI-waarde overgenomen.
 - Indien géén overeenkomende WISKI-O tijdstip gevonden wordt in de analyse-output, betekent dit dat geen YSI-reeks beschikbaar is voor dit moment (omdat het loggeheugen volzet was, of omdat geen logbestand werd opgeslagen op de cluster), en blijft de WISKI-P-waarde dus behouden.

De gecorrigeerde WISKI-P reeks wordt dan weggeschreven naar een nieuw ZRX-bestand.

4.2 Methodologie YSI analyse

De analyse wordt per YSI-bestand apart uitgevoerd en heeft als basis twee dataframes: die van het YSI-bestand (kolommen timestamp, YSI_value), en die van de WISKI-O-reeks (kolommen timestamp, WISKI_O_value). Het programma loopt in twee fasen: In een eerste fase wordt een eerste overeenkomst gezocht op basis van correlatie-analyse. De tweede fase bestaat dan uit het chronologisch controleren van volgende tijdstippen, en, indien de gevonden koppeling niet meer correct is, het zoeken van een nieuwe koppeling.

4.2.1 Fase 1: zoeken van eerste 'match':

Aangezien de WISKI-O reeks voor het gehele jaar werd ingelezen, en de YSI-reeks slechts bestand per bestand (ongeveer 1,5 maanden), dienen allereerst beide dataframes op min of meer gelijke lengte geclipt te worden, met inbegrip van enkele uren vooraan en achteraan om te compenseren voor de verwachte tijdsverschuiving.

Vervolgens wordt van start gegaan met het detecteren van de initiële tijdsfout in het YSI-bestand (er wordt van uitgegaan dat de tijdsfout niet constant is over de periode van het YSI-bestand). Dit gebeurt als volgt:

- De WISKI-waarde die dichtst bij het eerste YSI-tijdstip ligt wordt geselecteerd. In een range van 20 punten voor en na dit punt wordt gezocht naar een WISKI-O waarde die gelijk is aan de YSI-waarde (rekening houdend met de vooraf bepaalde tolerantie: [0,25 NTU voor Turbiditeit, 0,01 °C voor temperatuur]). Indien géén gelijke waarde wordt gevonden, wordt verdergegaan met de volgende YSI-waarde.
- Indien wél een waarde wordt gevonden wordt, startende van deze waarde, een reeks van 100 waarden geselecteerd (500 indien de YSI-reeks over minutelijkse waarden beschikt, terwijl de WISKI-O-reeks slechts 5-minutelijkse waarden heeft: hierbij worden beide reeksen eerst samengevoegd op basis van gemeenschappelijke tijdstippen, zie code voor meer uitleg).
- Ook startende vanaf de YSI-waarde wordt een reeks van 100 waarden geselecteerd, en de correlatie tussen beide selecties wordt berekend en, samen met de geselecteerde tijdsverschuiving, opgeslagen in een array. Wanneer 500 correlaties berekend zijn, wordt de tijdsverschuiving die algemeen de hoogste correlatie had, geselecteerd. De eerste waarde waarbij deze tijdsverschuiving gevonden werd, wordt de basis van de correctie fase. Alle voorgaande YSI-tijdstappen worden weggeknipt, ook bij de WISKI-O reeks worden alle voorgaande stappen verwijderd.

4.2.2 Fase 2: zoeken van overeenkomsten voor de rest van de reeks

Voordat de correctie van start gegaan worden voor twee uitzonderlijke situaties maatregelen getroffen:

- De meeste reeksen werden begin 2011 minutelijks opgemeten in plaats van vijfminutelijks. Echter, de waarden werden slechts vijfminutelijks naar de databank doorgeseind. Dit wil zeggen dat er in de reeksen resulterende na fase 1 de WISKI-O-reeks 5 maal het aantal waarden van de YSI reeks bevat. Voor elke WISKI waarde wordt de dichtstbijzijnde YSI waarde geselecteerd (mits toepassing tijdscorrectie).
- Sommige WISKI reeksen bevatten lange periodes met 1 constante waarde. De YSI-bestanden voor deze periodes worden aan het einde van het script behandeld. Aangezien geen correlatie gevonden kan worden en dus ook geen tijdsverschuiving berekend kan worden, wordt de gemiddelde tijdsverschuiving toegepast uit alle voorgaande geanalyseerde YSI-bestanden van de post/jaar combinatie.

Vervolgens worden de reeksen naast elkaar gelegd en chronologisch doorlopen. De YSI waarde wordt dan vergeleken met de overeenkomend WISKI waarde (mits toepassing huidige tijdsverschuiving). Indien deze overeenkomen binnen de bepaalde tolerantiegrens wordt overgegaan naar de volgende waarde. Indien ze niet overeenkomen wordt zijn er twee mogelijkheden: ofwel is de tijdsverschuiving nog geldig maar wijkt de WISKI waarde teveel af van de YSI waarde, ofwel is er een nieuwe tijdsverschuiving. Er wordt dus gezocht naar een nieuwe tijdsverschuiving door, zoals in fase 2, voor een range van 20 waarden voor en na het overeenkomende WISKI punt een reeks te selecteren en daarin te zoeken naar een waarde binnen de tolerantiegrens. Indien geen waarde gevonden wordt verdergegaan met de volgende YSI waarde. Indien wel zo'n waarde gevonden wordt, kan deze nieuwe tijdsverschuiving echter niet zomaar worden aangenomen. Er wordt voor de volgende 100 tijdstappen een correlatie berekend. Ook voor de huidige tijdsverschuiving wordt een correlatie van de volgende 100 punten berekend. Indien de nieuwe tijdsverschuiving beter correleert dan de huidige wordt deze aanvaard, en er wordt verder gerekend met de nieuwe tijdsverschuiving. Indien beide correlaties zeer laag zijn, wil dit zeggen dat er iets grondig mis is met de reeksen en wordt overgegaan naar het volgende YSI-bestand. Indien de huidige tijdsverschuiving beter blijft dan de nieuwe wordt verder gezocht naar een nieuwe tijdsverschuiving. Indien voor een van de volgende punten wel terug een match gevonden wordt blijft de huidige tijdsverschuiving behouden.

Op het einde van elk YSI-bestand blijven twee reeksen over, 1 YSI en 1 WISKI met telkens evenveel waarden. Deze reeksen worden naar een csv bestand weggeschreven en kunnen gebruikt worden om de nieuwe zrx reeksen aan te maken.

4.3 Toekomstige correcties: 2016

Volgende stappen dienen doorlopen te worden voor een correctie van toekomstige meetwaarden.

4.3.1 Export van WISKI-O reeksen

- Per te corrigeren post dient een folder 2016 aangemaakt te worden op de locatie:
 - P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\WISKI\Post
- In deze folder dienen de volgende folders aangemaakt te worden (naamgeving dient behouden te worden): Cond/ODO/ODOsat/ORP/pH/Press/Temp/Turbid.
- De export van WISKI-O reeks dient in deze folders ondergebracht te worden per parameter (enkel het jaar 2016, indien de volledige tijdsreeks wordt geëxporteerd kan het importeren in het script zeer lang duren).

4.3.2 YSI – interne log bestanden

- De interne log bestanden van 2016 worden bijgeplaatst op de locatie:
 - P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\YSI_bestand\Post
- Indien enkel het binaire *.dat formaat beschikbaar is, dient eerst mbv de software EcoWatch Lite een omzetting te gebeuren. Onder Parameters -> My template dienen volgende parameters geselecteerd te worden (in volgorde): Date D/M/Y – Temp C – SpCond mS – Cond mS – Press psia – pH – pH mV – Orp mV – Turbid+ NTU – ODOsat % - ODO mg/L.
- Het bestand kan dan geëxporteerd worden naar csv formaat. Een laatste omzetting naar *.txt dient te gebeuren waarbij ook dubbele aanhalingstekens worden verwijderd en decimale komma's naar punten worden omgezet. Dit kan aan de hand van het script replace.py (te vinden onder P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\bin). Plaats dit script in de map waar de csv bestanden zich bevinden en voor uit door adhv **CMD** commando: *python replace.py*. Plaats de resulterende *.txt bestanden onder P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\YSI_bestand\Post.

4.3.3 Runnen master_YSI_frombatch.py

- Open een command line window in de folder P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\bin (shift+rechtsklik -> Opdrachtvenster hier openen).
- Voer 1 van volgende commando's in:
 - 1 post, 1 parameter: *python master_YSI_frombatch.py **post Turbid o** 2016*
 - 1 post, alle parameters: *python master_YSI_frombatch.py **post all o** 2016*
 - Verschillende posten, all parameters:
*For %p in (Melle, Aarschot, Grobbendonk, Kruibeke) do start python master_YSI_frombatch.py **%p all o** 2016*
- De gecorrigeerde zrx data komt terecht in P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing**Post**\corrected_zrx\fullcomment\2016

5 YSI-data correctie

De voor het Moneos rapport (2016) prioritaire posten werden voor alle beschikbare jaren gecorrigeerd. Voor de overige posten werden eerst plots gemaakt die een aanduiding geven van de aanwezige afwijkingen tussen de interne log bestanden en de WISKI-O reeks. Deze aanduiding is gebaseerd op een eenvoudige correlatie tussen YSI en WISKI reeks, en op interpolatie van de WISKI reeks naar de YSI tijdstappen, waardoor bijkomend kleine afwijkingen geïntroduceerd worden. Bepaalde periodes met grote en frequente afwijkingen werden hieruit geselecteerd en gecorrigeerd. Hieronder volgt een olijsting van gecorrigeerde periodes en afwijkingen. Voor de posten waarvoor deze analyse werd uitgevoerd zijn de locaties van de html-plots hieronder weergegeven (zie Tabel 2).

Tabel 2: Locaties van analyseplots.

Post	Locatie plots
Aarschot	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Aarschot\diff_plots
Destelbergen	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Destelbergen\diff_plots
Gavere	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Gavere\diff_plots
Grobbendonk	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Grobbendonk\diff_plots
Itegem	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Itegem\diff_plots
Lippenbroek_in	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Lippenbroek_in\diff_plots
Lippenbroek_uit	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Lippenbroek_uit\diff_plots
Wondelgem	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Wondelgem\diff_plots
Zulte	P:\12_076-validsedfys\Project bestek IMDC WL_2015_26\Uitvoering\YSI\processing\Zulte\diff_plots

In deze sectie wordt voor elk van de gecorrigeerde posten een samenvatting van de correctieresultaten gegeven. Er worden steeds twee tabellen weergegeven. De eerste tabel is gebaseerd op de output van de analyse-fase, het csv document waarin elke overeenkomst tussen YSI interne log en WISKI-O wordt weergegeven voor de parameter Turbiditeit. De tabel geeft de fractie weer van overeenkomsten die gelijke waardes bevatten ($\Delta = 0$), de fractie van overeenkomsten die afwijken met meer dan 0,25 NTU ($\Delta > 0,25$ NTU), en de fractie van overeenkomsten die afwijken met meer dan de tolerantiegrens die werd vastgesteld tijdens de correctie van de P-reeks ($\Delta > \text{tol}$). Als laatste wordt ook de fractie van waarden gegeven die als 'potentiële stotter' werden geïdentificeerd.

Een tweede tabel geeft voor elke parameter/tijd combinatie de fractie van waarden weer waarvoor de status op 'unchecked' werd geplaatst: dit is het verschil in aantal waarden met status 0 na de correctie ten opzichte van het aantal waarden met status 0 voor de correctie, gedeeld door het totale aantal waarden.

Bij de interpretatie van deze resultaten dienen een aantal zaken in het achterhoofd gehouden te worden: de fractie van waarden die op 'unchecked' geplaatst werden bij correctie van de P-reeks is niet noodzakelijk

gelijk aan de som van de fractie afwijkingen groter dan de tolerantie en de fractie potentiële stoters. Veelal is de fractie in de tweede tabel lager dan de som van de fracties in de eerste tabel, omdat een aantal waarden in de P-reeks tijdens de validatie reeds op status 'unchecked' geplaatst werden, en dus niet meegeteld worden in de fractie in de tweede tabel.

In sommige gevallen is de fractie in de tweede tabel hoger dan de som van de fracties in de tweede tabel. Dit komt dan doordat voor een bepaalde periode geen YSI interne log bestand beschikbaar was (omdat het geheugen van de logger volgelopen was). Hierdoor werden meer waarden op status 'unchecked' geplaatst tijdens de correctiefase dan gedetecteerd in de analysefase.

Een laatste mogelijkheid is dat de tweede tabel nulwaardes geeft, terwijl in de analysefase te corrigeren waarden werden gevonden. Dit kan komen doordat sommige reeksen nog niet gevalideerd werden (Kruibeke) en hierdoor alle waarden nog status 'unchecked' hadden.

5.1 Prioritaire posten

5.1.1 Melle

Melle werd gecorrigeerd voor de jaren 2015, 2014, 2013, 2012, 2011. Tabel 3 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 3: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > \text{tol}$	Stotter
2015	83420	10%	16%	1,18	12%	0%
2014	64551	2%	5%	2,40	1%	1%
2013	87251	3%	17%	1,53	15%	0%
2012	96371	2%	10%	1,36	8%	1%
2011	73767	1%	49%	2,60	34%	0%

Tabel 4 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 4: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

Melle	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	11%	4%	12%	1%	1%	2%	7%	7%
2014	3%	2%	4%	2%	2%	2%	2%	2%
2013	13%	12%	14%	1%	1%	1%	6%	4%
2012	7%	3%	7%	0%	0%	0%	2%	0%
2011	5%	3%	4%	4%	2%	2%	3%	6%

5.1.2 Dendermonde

Dendermonde werd gecorrigeerd voor de jaren 2015, 2014, 2013, 2012. Van 2011 zijn geen interne log-bestanden beschikbaar. Tabel 5 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 5: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	63696	5%	14%	3,76	2%	3%
2014	74384	5%	10%	0,53	8%	0%
2013	83492	7%	12%	0,48	11%	1%
2012	98352	7%	3%	2,59	1%	0%

Tabel 6 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 6: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	5%	3%	4%	3%	3%	3%	4%	29%
2014	9%	4%	8%	4%	4%	4%	6%	4%
2013	12%	8%	10%	2%	4%	3%	11%	6%
2012	1%	1%	2%	1%	2%	1%	1%	1%

5.1.3 Duffel

Duffel werd gecorrigeerd voor de jaren 2015, 2014, 2013, 2012, 2011. Tabel 7 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 7: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	82249	3%	4%	1,79	2%	0%
2014	84780	2%	13%	0,33	13%	0%
2013	97380	2%	1%	0,39	0%	0%
2012	101693	2%	6%	0,56	6%	0%
2011	69713	2%	26%	0,64	23%	0%

Tabel 8 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 8: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	3%	1%	2%	1%	1%	1%	1%	19%
2014	12%	5%	11%	1%	1%	2%	6%	2%
2013	1%	0%	0%	0%	0%	0%	0%	0%
2012	5%	2%	4%	1%	0%	0%	0%	1%
2011	13%	4%	9%	3%	2%	2%	3%	2%

5.1.4 Eppegem

Eppegem werd gecorrigeerd voor de jaren 2015, 2014, 2013, 2012, 2011. Tabel 9 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 9: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	88177	3%	66%	3,69	24%	0%
2014	70881	0%	88%	3,64	42%	0%
2013	99643	0%	88%	4,27	40%	0%
2012	94039	0%	94%	3,21	53%	0%
2011	22555	0%	88%	2,77	45%	0%

Tabel 10 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 10: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	17%	16%	15%	12%	24%	12%	15%	16%
2014	23%	13%	4%	2%	2%	2%	6%	5%
2013	24%	19%	3%	0%	0%	0%	2%	5%
2012	48%	19%	3%	0%	0%	1%	15%	3%
2011	14%	9%	5%	4%	4%	4%	6%	5%

5.2 Niet-prioritaire posten

5.2.1 Kruikeke

Kruikeke werd enkel gecorrigeerd voor het jaar 2015, daar het station pas actief werd in april 2015. Tabel 11 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 11: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	18006	4%	72%	4,11	6%	0%

Tabel 12 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks). Hierbij dient opgemerkt te worden dat in de WISKI-O reeks van Kruikeke alle waarden al op 'unchecked' stonden, wat het nulresultaat verklaard.

Tabel 12: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat
2015	0%	0%	0%	0%	0%	0%	0%

5.2.2 Aarschot

Aarschot werd gecorrigeerd voor de jaren 2015,2014,2013,2012,2011. Tabel 13 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 13: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	80904	8%	5%	2,05	1%	0%
2014	70419	9%	15%	3,31	6%	0%
2013	93689	5%	15%	2,42	7%	0%
2012	103700	5%	14%	1,94	7%	0%
2011	63651	8%	13%	3,69	4%	2%

Tabel 4 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 14: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	4%	4%	4%	4%	4%	4%	4%	30%
2014	4%	3%	4%	2%	3%	3%	4%	5%
2013	6%	4%	2%	1%	1%	1%	2%	2%
2012	7%	1%	3%	0%	0%	0%	0%	0%
2011	9%	6%	6%	6%	6%	6%	6%	6%

5.2.3 Destelbergen

Destelbergen werd enkel gecorrigeerd voor 2015, na visuele controle van de afwijkingen in alle jaren. Tabel 15 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 15: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	84998	14%	17%	3,19	11%	0%

Tabel 16 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 16: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat
2015	10%	7%	5%	0%	4%	0%	0%

5.2.4 Gavere

Gavere werd na visuele controle van de afwijkingen over de jaren, enkel gecorrigeerd voor het jaar 2014. Tabel 17 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 17: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2014	53591	4%	2%	2,22	0%	0%

Tabel 18 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 18: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2014	4%	4%	4%	4%	4%	4%	4%	4%

5.2.5 Grobbendonk

Grobbendonk werd gecorrigeerd voor de jaren 2015,2014,2013,2012,2011. Tabel 19 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 19: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	73048	6%	5%	2,03	1%	0%
2014	77466	8%	7%	3,47	1%	0%
2013	73013	5%	11%	3,94	2%	1%
2012	105112	6%	10%	2,51	3%	0%
2011	60332	8%	24%	3,98	3%	1%

Tabel 20 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 20: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	0%	0%	0%	0%	0%	0%	0%	34%
2014	5%	4%	4%	4%	3%	4%	4%	4%
2013	2%	1%	0%	0%	0%	0%	0%	1%
2012	2%	1%	0%	0%	0%	0%	0%	1%
2011	6%	4%	3%	3%	1%	3%	3%	3%

5.2.6 Itegem

Voor station Itegem werd na visuele controle beslist geen correctie door te voeren.

5.2.7 Lippenbroek_in

Station Lippenbroek_in werd nog niet gecorrigeerd.

5.2.8 Lippenbroek_uit

Station Lippenbroek_uit werd nog niet gecorrigeerd.

5.2.9 Wondelgem

Wondelgem werd na visuele controle van de afwijkingen gecorrigeerd voor de jaren 2012,2011. Tabel 21 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 21: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2012	57164	5%	21%	3,99	5%	1%
2011	49392	8%	18%	3,97	3%	1%

Tabel 22 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 22: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2012	5%	2%	1%	0%	1%	0%	0%	0%
2011	5%	17%	3%	2%	2%	2%	3%	4%

5.2.10 Zulte

Zulte werd gecorrigeerd voor de jaren 2015,2014,2013,2012,2011. Tabel 23 geeft de afwijkingen tussen YSI interne log en de WISKI-O reeks aan.

Tabel 23: Afwijkingen tussen interne log en WISKI reeks.

	Aantal	$\Delta = 0$	$\Delta > 0,25$ NTU	Tolerantie	$\Delta > Tol$	Stotter
2015	78717	11%	18%	1,17	14%	0%
2014	52436	2%	11%	2,39	7%	0%
2013	86421	4%	6%	3,85	1%	0%
2012	99630	3%	7%	1,19	4%	0%
2011	46341	3%	25%	4,01	9%	0%

Tabel 24 bevat voor elk van de parameters de fractie van waarden waarvan de status bijkomend op 'unchecked' geplaatst werden (dus verschil van aantal 'unchecked' waarden gedeeld door aantal waarden in de O reeks).

Tabel 24: Fractie van WISKI-P reeks met status op 'unchecked' geplaatst.

	Turbid	Cond	Press	pH	Temp	ODO	ODOsat	ORP
2015	0%	0%	0%	0%	0%	0%	0%	35%
2014	7%	2%	2%	2%	2%	2%	3%	2%
2013	1%	0%	0%	0%	0%	0%	0%	0%
2012	4%	2%	4%	0%	3%	1%	4%	1%
2011	9%	3%	2%	1%	17%	1%	3%	2%

5.2.11 Lummen

Voor station Lummen zijn geen interne-log bestanden beschikbaar voor de periode 2015-2011.

5.2.12 Halen

Voor station Halen zijn geen interne-log bestanden beschikbaar voor de periode 2015-2011.

DEPARTEMENT **MOBILITEIT & OPENBARE WERKEN**
Waterbouwkundig Laboratorium

Berchemlei 115, 2140 Antwerpen

T +32 (0)3 224 60 35

F +32 (0)3 224 60 36

waterbouwkundiglabo@vlaanderen.be

www.waterbouwkundiglaboratorium.be