

Plain Urban Water Transparency Prediction

Subjects: Environmental Studies

Contributor: Jingxiang Shu

Water transparency is commonly used to indicate the combined effect of hydrodynamics and the aquatic environment on water quality throughout a river network. However, how water transparency responds to these indicators still needs to be explored, especially their complicated nonlinear relationship; thus, this study represents an analysis of the Suzhou civil river network. Using an artificial neural network (ANN) hydrological model and a multiple linear model (MLR) with in-situ data between 2013–2019, we investigated the Suzhou River's sensitivity to the six factors and water transparency, which including flow velocity and data from five categories of water-quality monitoring data: total suspended matter (TSS), water temperature (TE), dissolved oxygen (DO), chlorophyll (Chl) and chemical oxygen demand (COD). The results suggest that the ANN model can achieve better performance than the MLR model. Furthermore, results also show a well-established correlation between enhanced hydrodynamics and improved water transparency when the flow velocity ranged from 0.22 to 0.45 m/s. Overall, COD is a vital factor for the SD prediction because including the COD can see a notable improvement in the ANN model (with a correlation coefficient of 0.918). This study demonstrates that the ANN model with hydrodynamic and water quality parameters can achieve a better prediction of water transparency than other discussed models for a coastal plain urban river network.

Keywords: hydrodynamic ; water transparency ; plain urban river network ; correlation coefficient ; machine learning

1. Introduction

Urbanization has accelerated the degradation of urban aquatic ecosystems, and the associated ecological issues have been recognized in China ^{[1][2][3]}. To mitigate these effects and improve urban river health, large-scale water-clearing regulations have been instituted for plain urban river networks, such as the one that empties into the Yangtze River Delta ^{[4][5]}. Water transparency is a commonly used indicator of water quality ^[6] because it incorporates physical, chemical, and biological processes ^{[7][8]} (e.g., flow rate, nutrient cycling, and phytoplankton photosynthesis, respectively ^{[9][10][11]}). Moreover, it is a key indicator for measuring the effect of ecological restoration ^{[12][13]} and is widely valued in China when constructing urban water environments ^{[14][15][16][17][18]}.

Additionally, studies have shown that the three abovementioned processes also relate to geographical location ^{[19][20][21]}. For instance, water transparency is dominated by dynamic conditions associated with wind, waves, and human activities in shallow areas such as the Yangtze Delta. Other environmental water quality indicators include chlorophyll-a (Chl-a) ^[22] ^[23]; nutritional status ^[24]; total phosphorus ^[25]; sediment resuspension, transportation and settlement; and total suspended substance ^[26]. For a long time, the Yangtze River was diverted to bring water to the cities on the coastal plain, but this method was once considered to be a waste of money and labour. Therefore, an investigation into how hydrodynamic and hydro-environmental factors affected water transparency was begun. This improved understanding of the large-scale water clearing regulation can provide a theoretical basis for a better evaluation of water diversion projects.

The Secchi disk depth, or Secchi depth (SD), is a simple, traditional measure of water transparency ^{[27][28]}. A black-and-white disk is immersed vertically into water, and the visual depth is called the Secchi depth ^[29], indicating the extent of the water transparency. Although the Secchi disk is a powerful tool, its main disadvantage is discrete spatial-temporal and asynchronous observations ^{[30][31]} for large areas. Additionally, errors may occur due to a lack of visual acuity. Therefore, the traditional method might not be adequate to evaluate water quality for large river networks. The Secchi method also consumes an abundance of labour and resources because of the complex branch system of a plain urban river network. As such, satellite sensors like China's CBERS-1 can provide high-quality water transparency measurements with high spatial-temporal resolution ^[32]. Colour sensors combined with remote sensors have been successfully applied to estuarine and thalassic water ^[21], but there are limited studies on its application to China's inland shallow areas.

Previous studies have shown that SD is a function of hydrodynamic factors like velocity and water level ^{[33][34]}. Further, hydrodynamic variations may lead to variations in hydro-environmental indicators ^[35], and variations in hydro-environmental indicators might further cause variations in SD. As such, only employing hydrodynamic indicators to predict

SD might be inadequate by themselves. For this reason, previous studies have explored the function of SD with various environmental parameters [32][36][37][38][39]. A partial list is given in **Table 1**). Over the past 50 years, several regression models have been developed, and in those models, the SD (its natural or decimal logarithm) was shown as a function of one or two parameters. Compared to the regression model, studies that employed and ANN on SD predictions is limited to date [40].

Table 1. Selected studies for different Secchi depth (SD) prediction models.

Authors	Model	Inputs	Output	R ²
Carlson (1977) [33]	$\text{Ln}(\text{SD}) = 2.040 - 0.68\text{Ln}(\text{Chl-a})$	Chl-a	SD	0.86
Carlson (1977) [33]	$\text{Ln}(\text{SD}) = 3.876 - 0.98\text{Ln}(\text{TP})$	TP	SD	-
Brezonik (1978) [35]	$\text{Log}(\text{SD}) = 0.63 - 0.55\text{Log}(\text{Chl-a})$	Chl-a	SD	0.75
Brezonik (1978) [35]	$\text{Log}(\text{SD}) = 0.48 - 0.72\text{Log}(\text{TUR})$	TURd	SD	0.53
Gikas et al. (2006) [36]	$\text{SD} = 0.52(\text{Chl-a}) - 0.05$	Chl-ae	SD	0.02h
Gikas et al. (2006) [36]	$\text{SD} = 0.85(\text{Chl-a}) - 0.22$	Chl-ae	SD	0.34i
Gikas et al. (2009) [37]	$\text{Log}(\text{SD}) = 5.32\text{Log}(\text{Chl-a}) - 0.38 + 2.11\text{Log}(\text{TSS}) - 0.16$	Chl-ae, TSS	SD	0.37
Gikas et al. (2009) [37]	$\text{Log}(\text{SD}) = 10.96\text{Log}(\text{TP}) - 0.54$	TPf	SD	0.35
Wu et al. (2009) [38]	$\text{Ln}(\text{SD}) = -0.712 + 0.093\text{WL} - 0.278\text{WV} [\text{WL} \geq 14.75]$	WL, WV	SD	0.72
Wu et al. (2009) [38]	$\text{Ln}(\text{SD}) = -19.887 + 1.393\text{WL} - 0.217\text{WV} [\text{WL} < 14.75]$	WL, WV	SD	0.72

Abbreviations: SD: Secchi depth in Centimeters (cm or feet); Chl-ae: chlorophyll-a ($\mu\text{g/L}$); TPf: total phosphorus ($\mu\text{g/L}$); WL: water level in meters; WV: wind velocity (m/s); 0.34i: model from 1998 to 1999.

Nevertheless, limited studies were using the ANN for SD prediction considering the three processes (physical, chemical, biological) impact on the plain urban river network. Inspired by big-data analysis and machine learning techniques, we attempted to develop a machine learning model based on remotely sensed SD and other hydrodynamic and environmental parameters from 2013 to 2019 to assess the response of SD to the large-scale water clearing regulation in the Yangtze Delta. Selected input candidates for the machine learning model include hydrodynamic condition index and water environmental factors: surface velocity (V), total suspended solids (TSS) concentration, dissolved oxygen (DO) concentration, near-surface chlorophyll (Chl) concentration, chemical oxygen demand (COD) concentration, and water temperature (TE).

The objectives of this study are:

- To evaluate the big data analysis and self-learning ability of the developed machine learning model in SD prediction for a plain urban river network with long-term field observations;
- To compare the SD prediction performance between a machine model and a regression model to provide a better prediction model and highlight suitable parameters.

2. Plain Urban Water Transparency Prediction

The ANN model and MLR model are compared based on their performances in (i) training, (ii) verification, and (iii) testing phases, with results summarized in **Table 2**. It appears that the ANN model is more accurate and consistent in different subsets since all the values of RMSE and MAE are similar, and all the correlation coefficients are also close to unity, and the performance of this model can be well demonstrated based on RMSE. It also shows that the ANN model results in a much higher value of the CC than the MLR model. The prediction results regarding the CC value during the verification phase showed an approximately 38.1% of improvement. In addition, the forecast results regarding the CC value during the test phase improved by approximately 36.9%. In some previous studies, the reported prediction of SD was not tested on the training dataset, which was due to the insufficient data size [41]. In this study, as we can see from **Table 2**, these are very encouraging results regarding the modelling of SD, and the results were fitted in all phases.

Table 2. Statistical parameters of the input data.

Parameters	Xmean	X _{max}	X _{min}	S _x	C _v	CC
SD (m)	0.558	1.216	0.147	0.439	0.692	1.000
V (m/s)	0.236	0.624	0.089	0.348	0.761	0.146
TSS (mg/L)	25.8	54.4	9.1	6.574	1.235	-0.497
TE (°C)	17.8	32.6	1.3	4.223	0.542	-0.134
DO (mg/L)	5.14	8.92	0.23	1.174	0.211	-0.085
COD (mg/L)	20.23	36.72	11.48	5.241	0.463	-0.363
Chl (µg/L)	14.96	24.67	6.72	3.621	0.283	-0.47

According to **Table 2**, the results show that during the verification phase, the ANN model shows a reasonable estimation of SD. Furthermore, an acceptable level can be observed using the model M1 and M3, and through the comparison of various statistical indices (CC, RMSE and MAE) expounds the performance of ANN models better than the MLR models, which demonstrates that the ANN method has the good advantage on predictive ability to acquire the SD of the plain urban river network. In the verification phase, using the ANN model, the best results are achieved using the M4 model. Therefore, in this comparison, the prediction performance of M4 is slightly better than that of M1 and M3. In the testing phase, as shown in **Table 3**, model M4 is always the best model, while for the MLR model, the M1 is the best model. In order to possess a good predictive ability, RMSE and MAE should be as low as possible, but for CC, the value of this parameter should be as high as possible.

Table 3. Statistical characteristics of different SD (m).

	Mean	Standard Deviation	Minimum	Maximum
Weekly average SD	0.558	1.79	0.26	0.89
Weekly maximum SD	0.803	1.95	0.544	1.103
Weekly minimum SD	0.321	1.70	0.179	0.566

Consequently, we can see that the inclusion of the two parameters (DO and TE) may not improve the performance of the model. Interestingly, besides TSS and Chl, the COD assumed major importance when included simultaneously as input to the model. As water quality parameters that affect the SD of the water body, when included with COD parameters, DO, and TE did not contribute significantly to model performance to predict the SD of the urban river network. As the most important environmental factors in water bodies, DO and TE mainly affect the degradation rate of pollutants in urban river networks. As the river network of plain cities in the Yangtze River Delta has undergone years of diversion and flow control, the water quality of the river network gradually improved and entered a steady state. Therefore, DO and TE, which affects the chemical process, are less sensitive to transparency. In contrast, COD is extremely difficult to degrade in urban river network water bodies and is closely related to TSS, which cause the SD to be more sensitive to COD in the plain urban river network. The final model selected to predict the SD of the urban river network in this study contained velocity, TSS, Chl and COD (M4). The inclusion of the DO and TE may not improve the model performance and even sometimes contribute to increasing the values of the error indices. Additionally, more suitable and fewer inputs will help to simplify the implementation and the calculations procedure, which improve the practicality of the model.

Finally, for a given regression model, the LR model exhibited a slightly lower RMSE for exponential correlation than power correlation, whereas the SVR model resulted in an opposite result. The result indicates that within a certain flow rate threshold, there is a positive correlation between transparency and flow rate, which reveals that the hydrodynamic factors of the plain river network have a significant impact on the water transparency and can be used as an effective parameter of the prediction model. Within the range of flow velocity 0.22–0.45 m/s, increased flow rate has a positive effect on SD. On the one hand, the improvement in hydrodynamics brought by water resources regulation does have a positive impact on the water environment of urban river networks. On the other hand, the method of improving the water environment through hydrodynamic regulation has an improved flow rate threshold, which means hydrodynamic control is not a once-and-for-all method.

Additionally, in the analysis of long-term SD changes and ANN model results, it is essential to consider the influence of flow velocity changes caused by water regulation. The comparison of the correlation coefficient shown in **Table 3** reveals

that flow velocity has a larger impact weight on the water transparency of urban river networks, and the absolute value of its correlation coefficient is ranked before dissolved oxygen and temperature.

3. Conclusions

In this study, an artificial neural networks model is proposed for estimating Secchi depth in a plain urban river network using long-term observed data. Through the comparison of results between the ANN model and MLR model, it reveals that the hydrodynamic parameters can be used as effective parameters for SD prediction models of the urban river network. Additionally, the impact of COD concentration on transparency is crucial in the river network due to the notable improvement with the inclusion of COD parameter as input in the model. The more accurate and more practical model of SD is the one with input parameters including flow velocity, TSS, COD, and Chl, and sensitivity ranks from high to low as TSS, Chl, COD and flow velocity.

In addition, ANN models perform better than the MLR models, which demonstrates the existence of a complex nonlinear relationship between SD and various parameters. The support vector machine was used to deduce the relationship between SD and hydrodynamic parameter, and a strong positive correlation was explored in this study when velocity range from 0.22 m/s to 0.45 m/s. Over 90% of data fall in the predicted intervals of SD for the method, reflecting the flow rate threshold of hydrodynamic regulation to improve water transparency in the urban river network.

References

1. Zhang, J.; Gilbert, D.; Gooday, A.J.; Levin, L.; Naqvi, S.W.A.; Middleburg, J.J.; Scranton, M.; Ekau, W.; Peña, A.; Dewitte, B.; et al. Nature and human-induced hypoxia and consequences for coastal areas: Synthesis and future development. *Biogeoscience* 2010, 7, 1443–1467.
2. The State Council of China. Action Plan for Prevention and Control of Water Pollution; People's Publishing House: Beijing, China, 2015.
3. Zhang, J.Y.; Li, Y.; Wang, X.J. Reconsideration on issues related to water ecological civilization construction in China. *China Water Resour.* 2016, 19, 8–11. (In Chinese)
4. Lu, X.Y. Impact of Water Diversion on the Growth of Dominant Eutrophic Algae in Lake Taihu; Nanjing Hydraulic Research Institute: Nanjing, China, 2013. (In Chinese)
5. Cui, G.B.; Chen, X.; Xiang, L.; Zhang, Q.; Xu, Q. Evaluation of water environment improvement by interconnected river network in plain area. *J. Hydraul. Eng.* 2017, 48, 1429–1437. (In Chinese)
6. Arnberger, A.; Eder, R. The Influence of green space on community attachment of urban and suburban residents. *Urban Urban Green* 2012, 11, 41–49.
7. Philips, E.J.; Aldridge, F.J.; Schelske, C.J.; Crisman, T.L. Relationships between light availability, chlorophyll a, and tripton in a large, shallow subtropical lake. *Limnol. Oceanogr.* 1995, 40, 416–421.
8. Kukushkin, A.S. Long-term seasonal variability of water transparency in the surface layer of the deep part of the Black Sea. *Russ. Meteorol. Hydrol.* 2014, 39, 178–186.
9. Abu Hanipah, A.H.; Guo, Z.R. Reaeration caused by intense boat traffic. *Asian, J. Water Environ. Pollut.* 2019, 16, 15–24.
10. Aarup, T. Transparency of the North Sea and Baltic Sea—a Secchi depth data mining study. *Oceanologia* 2002, 44, 323–337.
11. Blenckner, T.; Adrian, R.; Livingstone, D.M.; Jennings, E.; Weyhenmeyer, G.A.; George, D.G.; Jankowski, T.; Järvinen, M.; Aonghusa, C.N.; Noges, T.; et al. Large-scale climatic signatures in lakes across Europe: A meta-analysis. *Glob. Chang. Biol.* 2007, 13, 1314–1326.
12. Bricaud, A.; Morel, A. Light attenuation and scattering by phytoplanktonic cells: A theoretical modeling. *Appl. Opt.* 1986, 25, 571–580.
13. Bulkley, J.W.; Mathews, A.P. Water Quality Relationships in the Great Lakes: Analysis of a Survey Questionnaire; The University of Michigan: Ann Arbor, MI, USA, 1974; pp. 1–84.
14. Wang, C.; Wei, Z.; Zhang, L. Experimental study on improvement of water environment by water diversion in plain river networks. *J. Hohai Univ. Nat. Sci.* 2005, 33, 136–138. (In Chinese)
15. Xia, J.; Zhai, X.Y.; Zeng, S.; Zhang, Y. Systematic solutions and modeling on eco-water and its allocation applied to urban river restoration: Case study in Beijing, China. *Ecohydrol. Hydrobiol.* 2014, 14, 39–54.

16. Zhang, C.; Yu, Z.G.; Zeng, G.M.; Jiang, M.; Yang, Z.Z.; Cui, F.; Zhu, M.Y.; Shen, L.Q.; Hu, L. Effects of sediment geochemical properties on heavy metal bioavailability. *Environ. Int.* 2014, 73, 270–281.
17. Cai, J.; Wang, C.S.; Wang, W. Discussion on improving the water environment in the river course by diverting clean water and conducting better management. *China Water Resour.* 2011, 7, 39–41. (In Chinese)
18. Yang, Q.Q.; Wu, S.Q.; Wu, X.F.; Yang, Z.Z.; Cui, F.; Zhu, M.Y.; Shen, L.Q.; Hu, L. Effects of simulated water diversion on water quality and phytoplankton community in Meiliang bay. *J. Hydroecol.* 2015, 36, 42–49. (In Chinese)
19. Mukundan, R.; Pierson, D.C.; Schneiderman, E.M.; O'donnell, D.M.; Pradhanang, S.M.; Zion, M.S.; Matonse, A.H. Factors affecting storm event turbidity in a New York City water supply stream. *Catena* 2013, 107, 80–88.
20. Zhang, Q.H.; Yan, B.; Wai, O.W.H. Fine sediment carrying capacity of combined wave and current flows. *Int. J. Sediment Res.* 2009, 24, 425–438.
21. Chung, E.G.; Bombardelli, F.A.; Schladow, S.G. Modeling linkages between sediment resuspension and water quality in a shallow, eutrophic, wind-exposed lake. *Ecol. Model.* 2009, 220, 1251–1265.
22. Lorenzen, M.W. Use of chlorophyll-Secchi disk relationships. *Limnol. Oceanogr.* 1980, 25, 371–372.
23. Fleminglehtinen, V.; Laamanen, M. Long-Term changes in Secchi depth and the role of phytoplankton in explaining light attenuation in the Baltic Sea. *Estuar. Coast. Shelf Sci.* 2012, 102, 1–10.
24. Avigliano, E.; Schenone, N. Water quality in Atlantic rainforest mountain rivers (South America): Quality indices assessment, nutrients distribution, and consumption effect. *Environ. Sci. Pollut. Res.* 2016, 23, 15063–15075.
25. Johengen, T.H.; Biddanda, B.A.; Cotner, J.B. Stimulation of Lake Michigan Plankton Metabolism by Sediment Resuspension and River Runoff. *J. Great Lakes Res.* 2008, 34, 213–227.
26. Taillie, D.M.; O'Neil, J.M.; Dennison, W.C. Water quality gradients and trends in New York Harbor. *Reg. Stud. Mar. Sci.* 2020, 33, 100922.
27. Bessell-Browne, P.; Negri, A.P.; Fisher, R.; Clode, P.L.; Duckworth, A.; Jones, R. Impacts of turbidity on corals: The relative importance of light limitation and suspended sediments. *Mar. Pollut. Bull.* 2017, 117, 161–170.
28. Tyler, J.E. The Secchi disc. *Limnol. Oceanogr.* 1968, 8, 1–6.
29. Davies-Colley, R.J. Measuring water clarity with a black disc. *Limnol. Oceanogr.* 1988, 33, 616–623.
30. Preisendorfer, R.W. Secchi disk science: Visual optics of natural waters. *Limnol. Oceanogr.* 1986, 31, 909–926.
31. Larson, G.L.; Hoffman, R.L.; Hargreaves, B.R.; Collier, R.W. Predicting Secchi disk depth from average beam attenuation in a deep, Ultra-clear lake. *Hydrobiologia* 2007, 574, 141–148.
32. McClain, C.R. A decade of satellite ocean color observations. *Annu. Rev. Mar. Sci.* 2009, 1, 19–42.
33. Carlson, R.E. A trophic state index for lakes. *Limnol. Oceanogr.* 1977, 22, 361–369.
34. Kim, S.H.; Yang, C.S.; Ouchi, K. Spatio-Temporal patterns of Secchi depth in the waters around the Korean Peninsula using MODIS data. *Estuar. Coast. Shelf Sci.* 2015, 164, 172–182.
35. Brezonik, P.L. Effect of organic color and turbidity of secchi disk transparency. *J. Fish Res. Board Can.* 1978, 35, 1410–1416.
36. Gikas, G.D.; Yiannakopoulou, T.; Tsihrintzis, V.A. Water quality trends in a coastal lagoon impacted by non-point source pollution after implementation of protective measures. *Hydrobiologia* 2006, 563, 385–406.
37. Gikas, G.D.; Tsihrintzis, V.A.; Akrotos, C.S.; Haralambidis, G. Water quality trends in Polyphytos reservoir, Aliakmon River, Greece. *Environ. Monit. Assess.* 2009, 149, 163–181.
38. Wu, G.; Leeuw, J.D.; Liu, Y. Understanding seasonal water clarity dynamics of Lake Dahuchi from in situ and remote sensing data. *Water Resour. Manag.* 2009, 23, 1849–1861.
39. Yan, Z.; Ding, F.Y.; Qian, Y.; Pan, S.; Gai, Y.; Cheng, W.; Liu, X.; Tang, S. Variations of Water Transparency and Impact Factors in the Bohai and Yellow Seas from Satellite Observations. *Remote Sens.* 2021, 13, 514.
40. Xu, Y.X.; Wang, W.C.; Zeng, W.F.; Li, Y.; Lai, Q.; Yin, X.; Zhang, S. Simulation on improvement of water environment in plain river network by water diversion. *Water Resour. Prot.* 2018, 34, 70–76. (In Chinese)
41. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* 2011, 334, 1518–1524.

