DECIPHERING RIVER FLOW DATA TO DETERMINE RIVER PROPERTIES

Ahmed Moustafa, Ashraf El-Hamalawi and Andrew Wheatley

Civil and Building Engineering Department, Loughborough University, Loughborough, LE11 3TU, England (UK)

ABSTRACT

The advanced monitoring of water quality and performing a real time hazard analysis prior to entering Water Treatment Works (WTW) is very much a necessity nowadays in order to give warning of any contamination and avoid downtime of the WTW. Ultrasonic flow sensors are considered to be the most stable and reliable sensors to be used within the water and wastewater industry, as they are not easily affected by the weather conditions of the rivers' varying temperatures. Two case studies were looked at from two UK Rivers; one being the River Derwent in the East Midlands, Derbyshire, and the other being the River Severn in the West Midlands, Gloucestershire. Correlations between the flow rates of each river and the turbidity rates at a Water Treatment Works (WTW) near the gauging stations on the rivers were established. The river flow rates were then used to estimate the turbidity levels in the rivers prior to entering the WTW. The results showed that correlations existed between the rivers flow rates and its properties such as turbidity, colour and conductivity. The case studies highlighted the need for a new, intelligent and smart non-invasive type of measurement device to be introduced to enable the WTW to operate with optimum performance.

Keywords: Water Monitoring, Sensors, Turbidity, Flow Rate

1. INTRODUCTION

Water treatment plants must be able to produce a finished product of consistently high quality regardless of how great the demand might be. Like wastewater treatment, water treatment consists of a range of unit processes, usually used in series, which provides some design and operational flexibility. The treatment required by water prior to being delivered to consumers will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less. The most expensive operations in conventional treatment are sedimentation and filtration, while water softening can also be very expensive. Groundwater is generally much cleaner than surface water and so does not require the same degree of treatment, apart from aeration and disinfection before supply. Naturally occurring substances that may need to be reduced or removed in groundwater include iron, hardness (if > 300 mg/l as $CaCO3$) and Carbon dioxide.

Substances originating from humans are becoming increasingly common in groundwater and those requiring treatment include nitrates, pathogens and trace organics such as pesticides. Surface water requires more complex treatment due to its complex nature, although the quality of surface water can be very high (for example upland reservoir) (Geldreich, [1]; Reasoner, [2]).

In 2004, 375 sites were monitored for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales. Of these, 155 sites failed to comply with the Directive. However, over 90% of these 'failures' were due to insufficient sampling. These sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, problems at the laboratory, and sampling error. The quality of abstracted water generally improved since 1993. It was found that levels of colouration, nitrate and polycyclic aromatic hydrocarbons (PAHs) most commonly exceeded the Directive's standards in 2004 (Environment Agency, [3]).

Monitoring and assessing the quality of waters in streams, reservoirs, lakes, and estuaries are critical to improve water quality. Current techniques for measuring water quality involve insitu measurements and/or the collection of water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for monitoring, assessing, or managing water quality for an individual or multiple water bodies across the landscape. Remote sensing of indicators of water quality offers the potential of relatively inexpensive, frequent, and synoptic measurements using non-invasive sensors.

Major pollutants are suspended sediments (turbidity), pathogens, nutrients, metals, dissolved organic matter (DOM), pesticides, chlorophylls (algae, plants), temperature, and oils. Remote sensing applications to determine water quality are limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, [4]). Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls) which may have responded to an input of chemicals. These remote sensing techniques should improve our ability to monitor changes in the water topography and contents.

In this paper, correlations between the flow rates and various other properties such as turbidity, colour and conductivity for two UK Rivers were computed, and then used to predict these properties' levels at various other locations and points in time for the rivers and associated WTW/gauge stations. The case studies highlighted the need for a new, intelligent and smart non-invasive type of measurement device to be introduced to enable the WTW to operate with optimum performance. The decision to look at case studies was made in order to find out how the water industry is coping with the flux of data that is being generated by sensors and to see what sort of improvements and/or recommendations can be made based on these studies (Dolgonosov & Korchagin, [5]).

2. CASE STUDIES

The two sites examined were based on two completely different river sources. One site, which can be referred to as the St Mary's Bridge site, has its main water source from the River Derwent in East Midlands. The second site investigated, can be referred to as Saxons Lode, has it main water source from the River Severn in West Midlands. This combination of two different river sources and locations enabled us to compare the demands on the flow sensors used in most river water resources. The hypothesis was to identify whether a possible combination, with a minimal number of sensors, would make it possible to operate the processes of the WTW close to the two case studies more efficiently or whether new alternative sensors would be necessary. Alag et al. [6] reported that by combining information from many different sources, it would then be possible to decrease the uncertainty and ambiguity inherent in processing the information from a single sensor source. A large number of sensors measuring many different variables can collectively achieve a high level of accuracy and reliability, depending on their accuracy and reliability.

2.1 St Mary's Bridge

Using one of the River Derwent gauging stations, operated by the Environmental Agency, the daily mean flow rate was attained at St Mary's Bridge in Derbyshire County. The WTW being investigated has a river abstraction of around 100 Ml/day output from a lowland river with a flow of 30 to 100 m^3/s and is located south of the gauging station on the River Derwent.

The project was initially launched in order to improve the monitoring of the river water prior to entering the WTW to ensure the continuous operation at the WTW. Several problems had to be resolved, the most important of which were the shutdown of the WTW because of false alarms, and the time taken to analyse the water quality during which the water characteristics would have changed. This needed continuous intervention by the operating manager to resolve these problems. The operators covered a large area and a number of remote sites and so the frequency of false alarms was unmanageable and the sensors switched off.

The main finding of the analyses was a correlation between the daily mean flow rate and the turbidity measurements. Consequently, using a partial-least squares statistical method; it was found that turbidity levels in the rivers could be estimated using the current daily mean flow. The Partial least squares (PLS) method, developed in the 1960's by Herman Wold as an econometric technique, is a method for constructing predictive models when the factors are many and highly collinear. It is more commonly used in chemical engineering processes, but less so in civil engineering. PLS has been applied to monitoring and controlling industrial processes, where hundreds of controllable variables and dozens of outputs need to be processed (Dijkstra, [7]; Geladi & Kowalski, [8]; Stone & Brooks, [9]). Figure 1 illustrates the resulting correlation between the flow and the turbidity levels on the River Derwent, both actual and estimated, using the correlated river flow. Figure 2 depicts measured conductivity values against approximated values using the same statistical concept.

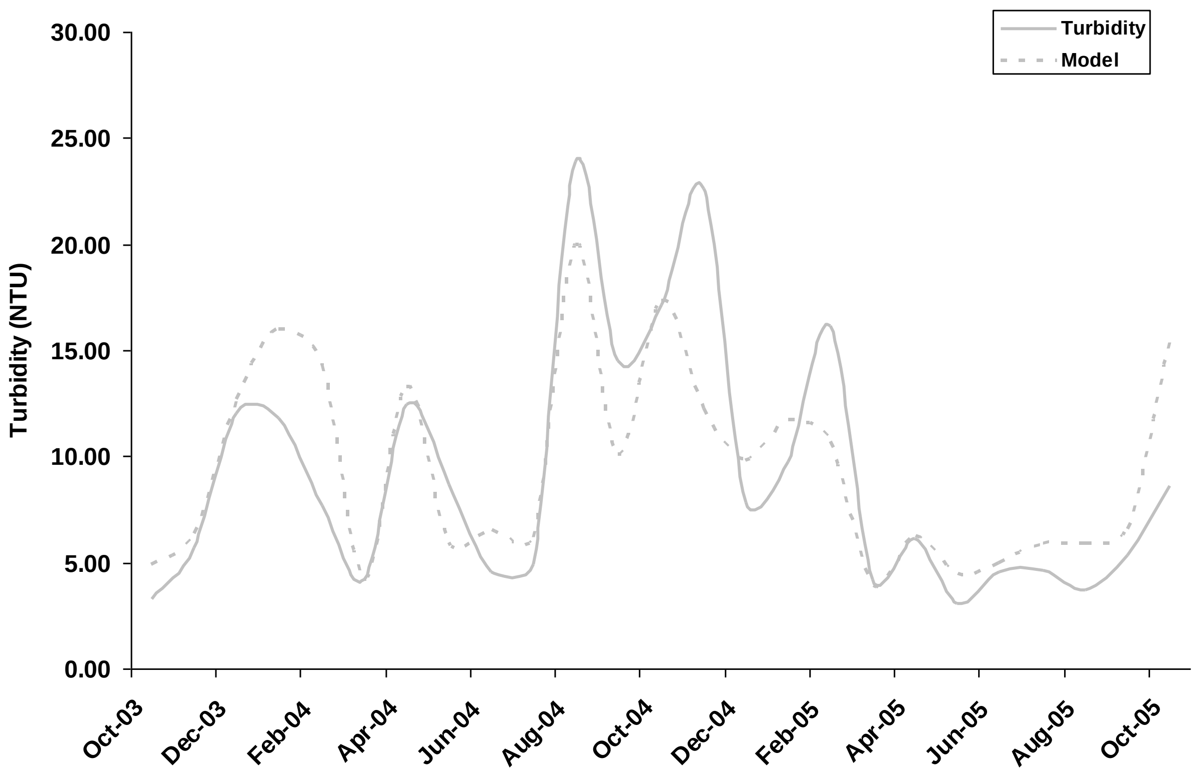

Figure 1: Measured Turbidity levels versus predicted levels extracted from the flow rate at St Mary's Bridge on the River Derwent.

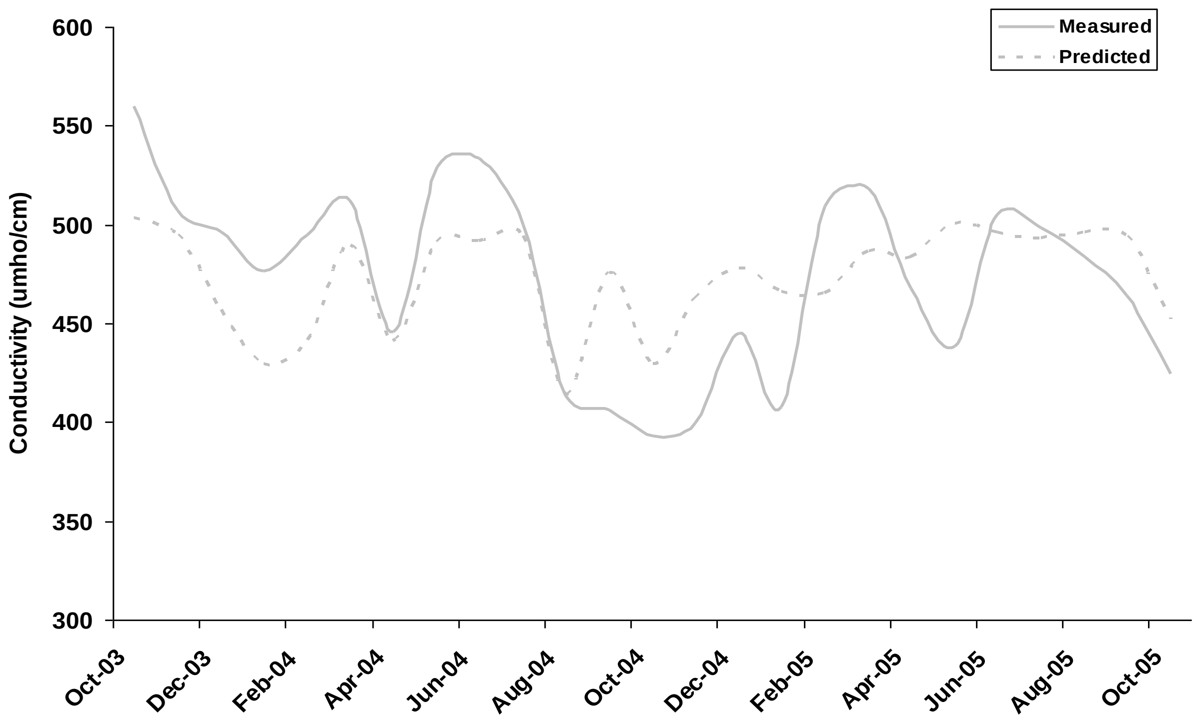

Figure 2: Measured Conductivity value against the predicted value extracted from the flow rate at St Mary's Bridge on the River Derwent.

Graphs 1 and 2 show a strong coherence and indicated the possible use of a very simple technique to predict the river water constituents and therefore cross check sensors output. It could help cut the maintenance cost and improve reliability by discarding spurious data under normal operational conditions. The analyses also indicate that it must also be possible to improve the quality of sensor information by using this type of expert system where a strong correlation can be established. PLS regression is an extension of the multiple linear regression models (e.g., Multiple Regression or General Stepwise Regression). In its simplest form, a linear model specifies the (linear) relationship between a dependent (response) variable Y, and a set of predictor variables, the X's, so that

$$[Y] = b_0 + b_1[X_1] + b_2[X_2] + ... + bp[Xp] \quad (1)$$

In this equation b_0 is the regression coefficient for the intercept and the bi values are the regression coefficients (for variables 1 through p) computed from the data. To put it in perspective, the dependent variable would be turbidity level, and the predictor variable is the daily mean flow rate.

PLS regression extends multiple linear regression without imposing the restrictions employed by discriminate analysis, principal components regression, and canonical correlation. In partial least squares regression, prediction functions are represented by factors extracted from the Y'XX'Y matrix. The number of such prediction functions that can be extracted typically will exceed the maximum of the number of Y and X variables. In short, PLS is the least restrictive of the various multivariate extensions of the multiple linear regression models. This flexibility allows it to be used in situations where the use of traditional multivariate methods is severely limited, such as when there are fewer observations than predictor variables. Furthermore, partial least squares regression can be used as an exploratory analysis tool to select suitable predictor variables and to identify outliers before classical linear regression.

2.2 Saxons Lode

A second case study was initiated to cross-validate the findings of the first case study. Attention was turned to the largest UK River, the River Severn. Again, the WTW concerned in this case was also downstream from the Saxons Lode gauging station. The fluctuations in the turbidity levels experienced in both case studies can be attributed to the fluctuations in the daily mean flow rate. This is a natural phenomenon and has more than one cause for its rise and fall. Some of the factors attributing to this would be rainfall, wind, flood seasons, seasonal temperature changes. Figure 3 illustrates the resulting correlation between the flow and the turbidity levels on the River Severn, both actual and estimated, using the correlated river flow.

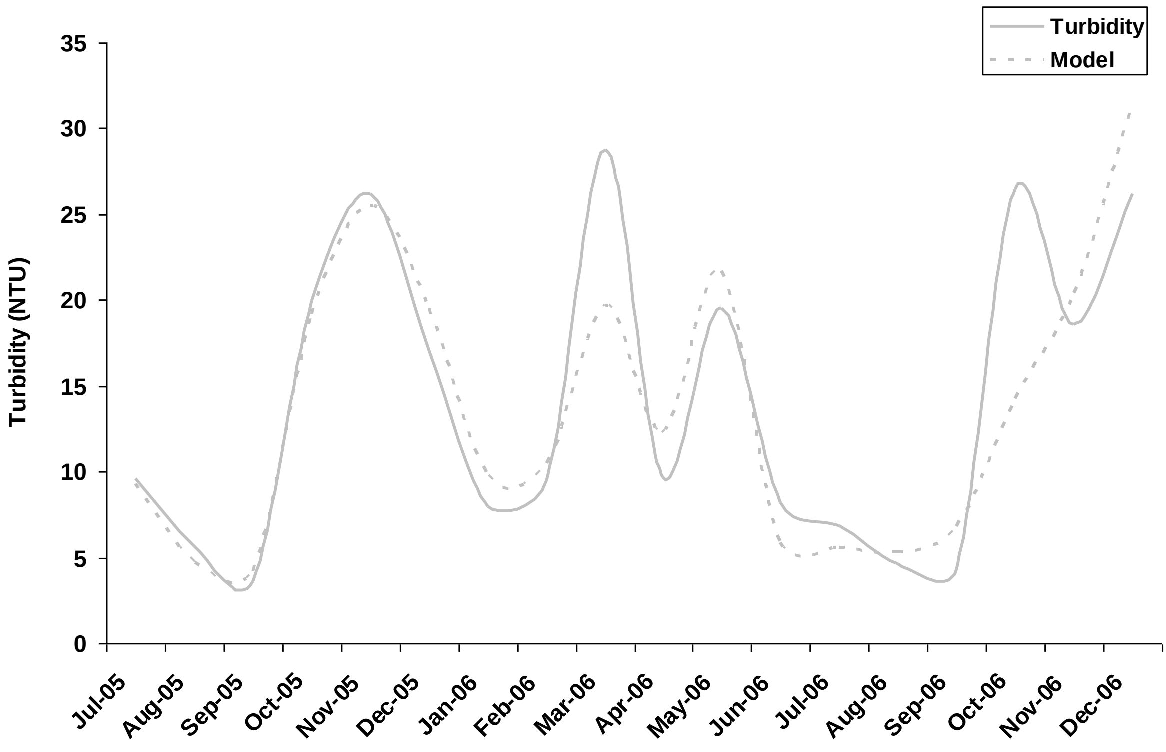

Figure 3: Measured Turbidity levels versus predicted levels extracted from the flow rate at Saxons Lode on the River Severn.

3. RESULTS AND DISCUSSION

For each of the WTW that were investigated; each site monitors Turbidity, Colour, pH, Conductivity, Ammonia and Temperature. Using this data, a partial least squares analysis was performed to try and correlate these data together. A very positive link was found between turbidity and colour, Fig. 4. This relation between turbidity and colour is logical; when turbidity is high then other parameters, e.g. colour, will also increase, since sediment load will contribute to the colouring and turbidity in the water. It was also found that the temperature was linked to conductivity readings, Fig. 5. This is as expected, since temperature would affect the solubility of contaminants, and therefore conductivity. The shift in time variation between the two graphs is attributed largely due to operator and sampling errors. For example, it is impractical to measure temperature of running water by sampling the water in the lab and not at the source.

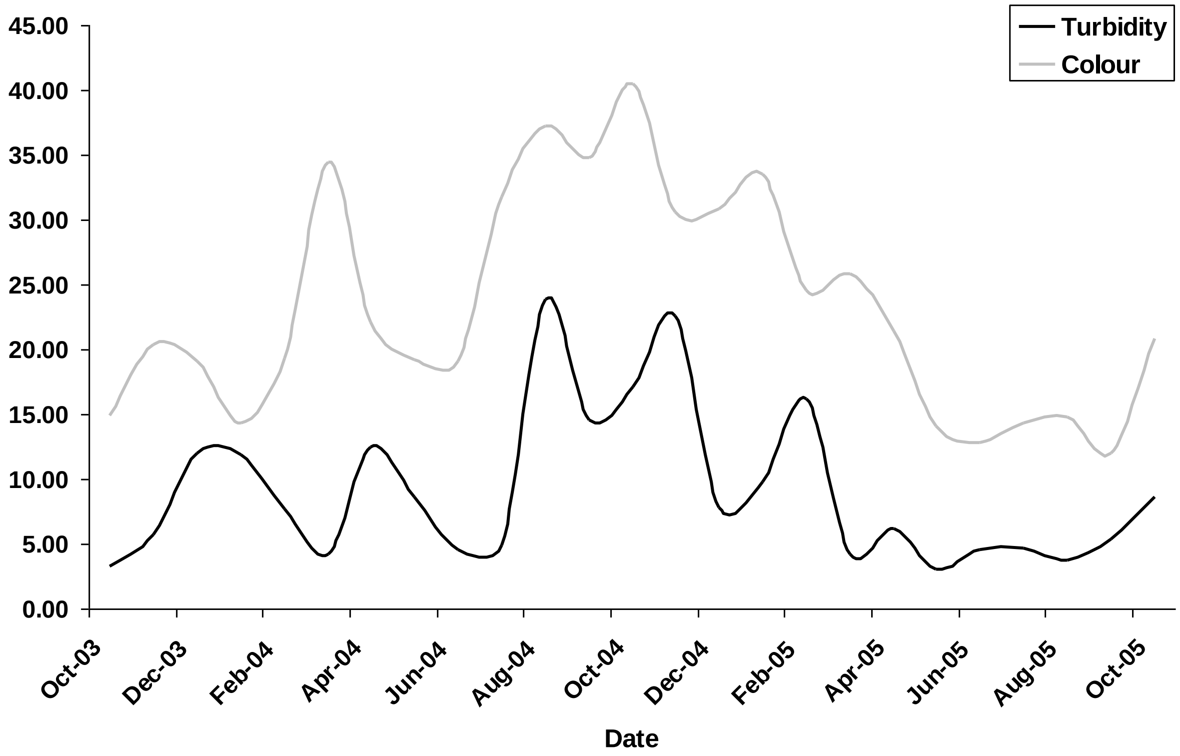

Figure 4: Graphical comparison between the Turbidity and Colour measurements at St Mary's Bridge on the River Derwent.

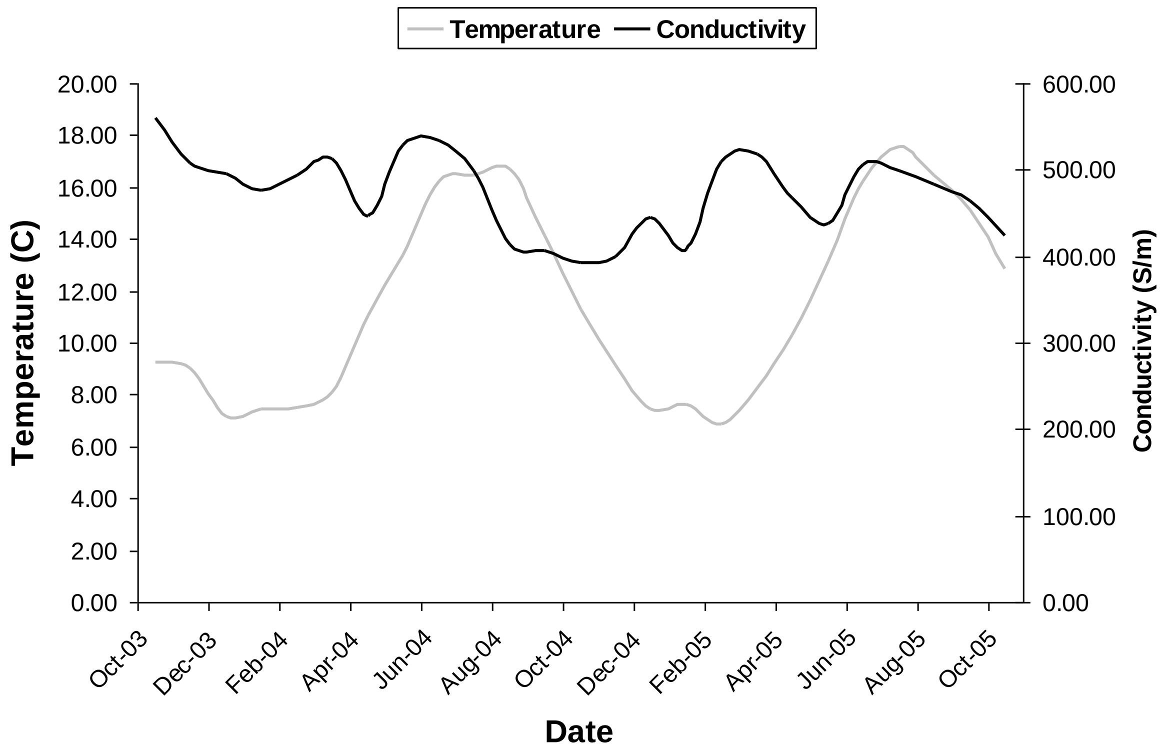

Figure 5: Graphical comparison between the Conductivity and Temperature measurements at St Mary's Bridge on the River Derwent.

From the analysis, it is also noticeable that there are relations between all the parameters including flow rate. This should not be surprising since an increase in flow is likely to increase the transport and erosion of materials. Figure 6 shows the complete data collected from the WTW with the daily mean flow rate. Following the derivation of the Turbidity level trend from the flow rate at St Mary's bridge, the data was fed back into the PLS regression, together with the flow data, and the predicted conductivity trend was calculated. The same process was repeated again by feeding the conductivity values back into the PLS method to derive the colour level. In simpler terminology, once the turbidity level is estimated, and using PLS, data fed back into PLS used conductivity as the dependent variable, while turbidity and flow account for the predictor variables. This process is repeated with colour as the dependent variable and conductivity is added to the predictor variables list.

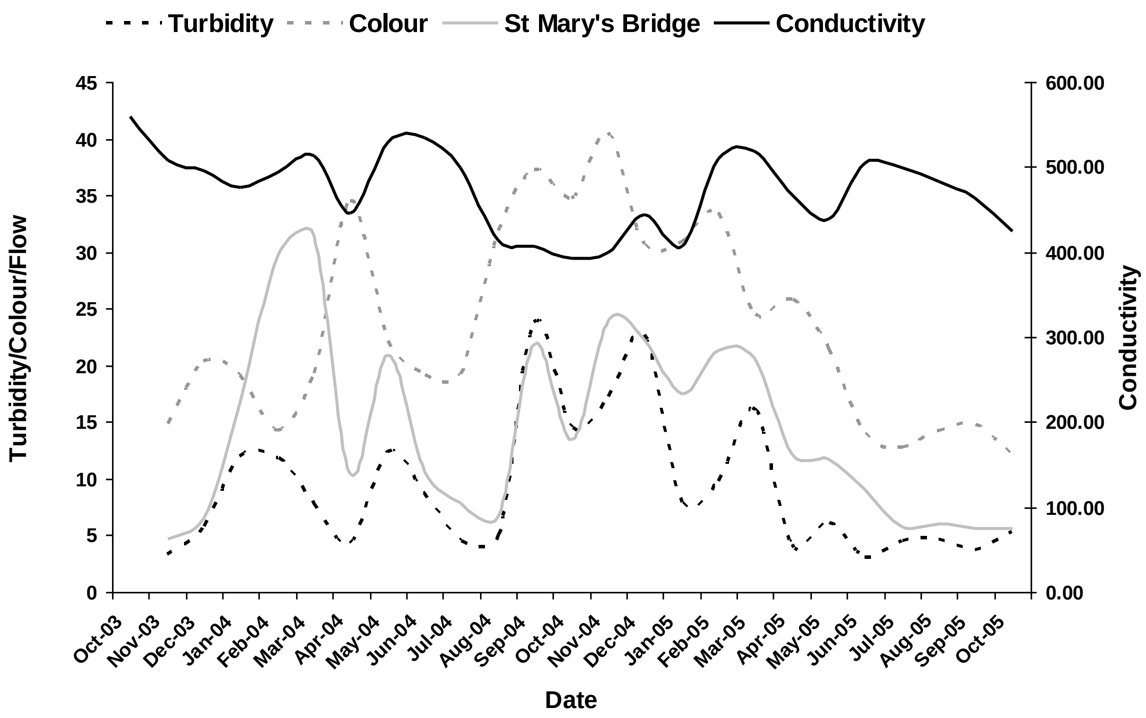

Figure 6: Graphical comparison between the three measured variables (turbidity, conductivity and colour) together with the daily mean flow rate at St Mary's Bridge, on the River Derwent, for two consecutive years.

It would be assumed that the calculated interpretations of the data for Turbidity, Conductivity and Colour could be more realistic of the real values in the River Derwent. It should however be stressed that these relationships should be developed as moving correlations over several years and they will not be appropriate coefficients for other WTW. These relationships were derived using the data supplied from one site. Different sites will have different data and correlations, but the same process could be used to save time and money, in order to increase the efficiency of predicting various parameters based on the flow going through WTWs. These correlation coefficients would vary according to flow, season (e.g. leaf fall), antecedent dry period and impoundment and other river engineering.

The interrelationship between sensors should be established to cross-check the other sensor performance. Data gathered also does not show how it benefits SCADA (Supervisory Control And Data Acquisition) or how it links with it. Finally, data from sensors are often disregarded; since operators intuitively find data from grab-wet samples more reliable.

4. CONCLUSION

Two case studies were looked at to try and examine the relationship between the rivers' flow rates and the rivers' constituents. Two river abstraction Water Treatment

Works were probed, one WTW on the River Derwent and the other on the River Severn. The findings presented for the two sites showed the lack of understanding of the rivers flow rates by the operators of the two WTW's and the potential use and benefits if it had been investigated much earlier. It also highlights the full potential of an alternative measuring and monitoring technique that can be implemented within the industry. Results emphasised the issue of backup monitoring and self adjusting automation processes that are needed within the industry. The study revealed that a relationship is needed to be found between the different types of sensors and/or measured parameters in order to cross-check the sensors' performance and be used as a guide of when maintenance procedures are needed.

ACKNOWLEDGEMENT

The authors would like to thank the Engineering and Physical Sciences Research Council (EPSRC) in the UK and Myriad Vision Ltd for their support and funding of this project. A special appreciation is also in place for Severn Trent Plc for the support given in accessing and gathering of the data for the two sites.

REFERENCES

[1] Geldreich, E.E., Microbial Quality of Water Supply in Distribution Systems, Lewis Publishers, Bocta Raton, Fl, USA, 1996.

[2] Reasoner, D., Pathogens in Drinking Water – Are there any new ones? US Environmental Protection Agency, Washington, D.C, USA, 1992.

[4] Ritchie, J.C. and Cooper, C.M. Remote sensing for determining water quality: Application to TMDL, pp. 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA, USA, 2001.

[3] Environment Agency, 2004. http://www.environment-agency.gov.uk/

[5] Dolgonosov, B. M., and Korchagin, K. A. Statistical assessment of relationships between water flow in a river and water turbidity in water intakes. Water Resources, vol. 32, no. 2, pp. 175-182.

[6] Alag S., Agogino A.M., Morjaria M. A methodology for intelligent sensor measurement, validation, fusion and fault detection for equipment monitoring and diagnostics. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 15, 307-320, 2001.

[7] Dijkstra, T. Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, 22, 67-90, 1983.

[8] Geladi, P., and Kowalski, B. Partial least squares regression: A tutorial. Analytical Chimica Acta, 185, 1-17, 1986.

[9] Stone, M. and Brooks, R. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B, 52(2), 237-269, 1990.