Water temperature modeling in the Garonne River (France)

K. Larnier(1,2), H. Roux(1,2), D. Dartus(1,2), O. Croze(1,2,3)

Received May 4, 2010 / Reçu le 4 mai 2010
Revised July 17, 2010 / Révisé le 17 juillet 2010
Accepted July 29, 2010 / Accepté le 29 juillet 2010

ABSTRACT

Stream water temperature is one of the most important parameters for water quality and ecosystem studies. Temperature can influence many chemical and biological processes and therefore impacts on the living conditions and distribution of aquatic ecosystems. Simplified models such as statistical models can be very useful for practitioners and water resource management. The present study assessed two statistical models – an equilibrium-based model and stochastic autoregressive model with exogenous inputs – in modeling daily mean water temperatures in the Garonne River from 1988 to 2005. The equilibrium temperature-based model is an approach where net heat flux at the water surface is expressed as a simpler form than in traditional deterministic models. The stochastic autoregressive model with exogenous inputs consists of decomposing the water temperature time series into a seasonal component and a short-term component (residual component). The seasonal component was modeled by Fourier series and residuals by a second-order autoregressive process (Markov chain) with use of short-term air temperatures as exogenous input. The models were calibrated using data of the first half of the period 1988–2005 and validated on the second half. Calibration of the models was done using temperatures above 20 °C only to ensure better prediction of high temperatures that are currently at stake for the aquatic conditions of the Garonne River, and particularly for freshwater migrating fishes such as Atlantic Salmon (Salmo salar L.). The results obtained for both approaches indicated that both models performed well with an average root mean square error for observed temperatures above 20 °C that varied on an annual basis from 0.55 °C to 1.72 °C on validation, and good predictions of temporal occurrences and durations of three temperature threshold crossings linked to the conditions of migration and survival of Atlantic Salmon.

RÉSUMÉ

Modélisation de la température de l’eau de la Garonne (France)

Mots-clés : température de l’eau,
INTRODUCTION

Stream water temperature is one of the most important parameters for water quality and ecosystem studies. Temperature can influence many chemical and biological processes – such as dissolved oxygen - and therefore impacts on the living conditions and distribution of aquatic ecosystems. The most obvious effects of temperature on aquatic organisms are on their survival and growth rate. For instance, conditions for Atlantic Salmon (*Salmo salar* L.) particularly depend on water temperature, and high temperatures widely disturb migration of this species (Decola, 1970; Chanseau *et al*., 1999; Fairchild *et al*., 1999; Swansburg *et al*., 2002). Moreover, temperatures above 24 °C may be considered lethal for this species (Alabaster, 1967; Elliott, 1991; Wilkie *et al*., 1997).

Understanding the thermal regime of watercourses is therefore very important for management of aquatic resources and fisheries. The thermal regime of a river is governed by many environmental processes (*e.g.* climatic conditions, topography, etc.) and by some human activities (Caissie, 2006; Webb *et al*., 2008). Knowledge of the main driving processes of the thermal regime is essential to understand the spatial and temporal variations in the water temperature better.

Ability to predict stream water temperature is also essential in conducting environmental studies and restoration plans. Problems of aquatic species related to water temperatures often happen for high summer temperatures. Predicting time occurrences and duration of the highest temperature periods could be very useful to trigger plans or technical tools to restore favorable aquatic conditions.

To predict water temperatures in streams, many models have been developed and used. These models are often categorized into two major groups: deterministic models and statistical models (Benyahya *et al*., 2007a). Deterministic models generally consider all relevant heat fluxes between the body of water and both the atmosphere and the bed material (Sinokrot and Stefan, 1994; Kim and Chapra, 1997). Such models are very useful for studies dealing with anthropogenic impacts and changes in inputs (changes in flow regime, watershed restriction, etc.) and distributed deterministic models can predict spatial variations along the watercourse. One major drawback of deterministic models is the need for large amounts of data and computational resources.
As a result, practitioners prefer to use simplified models such as statistical models. Numerous statistical models have been used in the literature. Among statistical models, regression models have been widely used (Stefan and Preud’homme, 1993; Pilgrim et al., 1998; Erickson and Stefan, 2000) and showed good results in predicting temperature at weekly or monthly time steps, relying on the relatively strong relation between air and water temperature on those time scales (Benyahya et al., 2007a). Most of those models were based on linear regression but in some cases, non-linear regression models have been used for a better description of the change in slope in the relation between air and water temperatures at both low and high air temperature (Mohseni et al., 1998). Although relatively efficient, regression models on a time scale shorter than weekly are more difficult to apply due to autocorrelations in the structure of water temperature time series. In these cases, stochastic models and non-parametric models such as Artificial Neural Networks (ANN) showed better results (Benyahya et al., 2007a; Chenard and Caissie, 2008).

Two particular statistical models – the equilibrium concept-based model and stochastic autoregressive models with exogenous inputs – have shown good efficiency when modeling daily mean water temperatures for large rivers (Caissie et al., 2005; Ahmadi-Nedushan et al., 2007; Benyahya et al., 2007b). These two models only use air temperature as a predictor and therefore the relation between air and water temperatures was to be assessed.

The objectives of this study were (a) to assess the influence of climatic conditions on water temperatures in the Garonne River in order to verify that air temperature is the main factor that has influenced the thermal regime of the Garonne River for the past two decades, and (b) to assess the efficiency of two statistical models to predict daily mean water temperatures, and particularly the high summer peaks that are currently an issue for aquatic ecology.

**METHODOLOGY**

**> TRENDS AND CORRELATIONS**

The first step of our study was to analyze trends in the evolution of stream water temperatures, and hydraulic and climatic parameters to determine parameters potentially related to the thermal regime evolution of the Garonne River. Analyses of trends in descriptive statistics of the time series (annual percentiles, annual and seasonal averages) were performed and related significances were assessed using the non-parametric Spearman rank correlation test. Afterwards, correlation analyses were performed between water temperatures and parameter statistics that presented significant trends.

**> WATER TEMPERATURE MODELS**

**Equilibrium concept**

Deterministic models in previous studies (Raphael, 1962; Sinokrot and Stefan, 1984; Morin and Couillard, 1990) have established the relevant energy components of heat exchange in rivers. The one-dimensional law of conservation of energy for vertically well-mixed streams is expressed as follows:

\[
\frac{\partial T_w}{\partial t} + u \frac{\partial T_w}{\partial x} - \frac{1}{A} \frac{\partial}{\partial x} \left( A \cdot D_L \frac{\partial T_w}{\partial x} \right) = \frac{B}{\rho \cdot C_w \cdot A} S_t + \frac{P}{\rho \cdot C_w \cdot A} S_{\text{bed}}
\]

where \( T_w \) is the water temperature (°C), \( t \) the time (day), \( x \) the longitudinal distance downstream (m), \( u \) the mean water velocity (m s\(^{-1}\)), \( A \) the cross-sectional area (m\(^2\)), \( D_L \) the longitudinal diffusive coefficient in direction of flow (m\(^2\) s\(^{-1}\)), \( B \) the width of the free surface, \( \rho \) the water density (1000 kg m\(^{-3}\)), \( C_w \) the specific heat of the water (4.85 × 10\(^{-2}\) W kg\(^{-1}\) °C\(^{-1}\)), \( S_t \) the net heat flux from the atmosphere to the river (W m\(^{-2}\)), \( P \) the wetted perimeter (m) and \( S_{\text{bed}} \) the heat flux with the streambed (W m\(^{-2}\)).
When dealing with water temperature on a daily basis or for longer time steps, the streambed heat flux can be neglected (Morin and Couillard, 1990; Sinokrot and Stefan, 1994). Furthermore, changes in temperatures along river reaches have been reported to be usually small compared with diurnal variation for river reaches with fairly uniform water temperature (Torgersen et al., 2001). In such cases, the diffusive and convective terms can be neglected in equation (1), which then can be simplified to the following form:

\[
\frac{\partial T_w}{\partial t} = \frac{B}{\rho \cdot C_w \cdot A} S_t
\]

(2)

where the parameters were defined previously. Equation (2) has been used in many studies to estimate water temperatures at specific locations of various streams using meteorological data (Marcotte and Duong, 1973; Morin and Couillard, 1990). Moreover, equation (2) can be used to estimate the upstream temperatures when conducting one-dimensional water temperature modeling (Sinokrot and Stefan, 1993).

The net heat flux \( S_t \) is a compound of net solar radiation, net long wave radiation, convection and evaporation and thus can be expressed using meteorological data only. Studies dealing with modeling the thermal regime of rivers have, however, shown that the net heat flux can be expressed in a simpler form using the equilibrium temperature concept (Edinger et al., 1968; Morin and Couillard, 1990). The equilibrium temperature stands for the water temperature leading to a null total heat flux \( (S_t(T_e) = 0 \) where \( T_e \) is the equilibrium temperature (°C)). Hence, the equilibrium temperature is a function of meteorological parameters. Methods for calculating the equilibrium temperature can be found in Mohseni and Stefan (1999) and Caissie et al. (2005). If such a temperature can be calculated, the net heat flux can be expressed using Newton’s law of cooling:

\[
S_t = K (T_e - T_w)
\]

(3)

where \( K \) is a thermal exchange coefficient \( (\text{W} \cdot \text{m}^{-2} \cdot \text{°C}^{-1}) \).

Using equation (3), equation (2) can be rewritten:

\[
\frac{\partial T_w}{\partial t} = \frac{B \cdot K}{\rho \cdot C_w \cdot A} (T_e - T_w).
\]

(4)

Influences of different physical and meteorological parameters can therefore be evaluated using the equilibrium temperature concept, as reported in Mohseni and Stefan (1999), which is one particular advantage of this concept.

Furthermore, although the equilibrium temperature is a function of many meteorological parameters, it can be reduced in temperate regions to a function of air temperature only. Indeed, strong linear association between the equilibrium temperature and air temperature can be postulated in such regions (Mohseni and Stefan, 1999). Using this hypothesis, equation (4) can be rewritten using air temperature (Caissie et al., 2005):

\[
\frac{\partial T_w}{\partial t} = K'(a_1 T_a + a_2 - T_w)
\]

\[
= \frac{K'}{h} a_1 T_a + a_2 - T_w
\]

(5)

where \( a_1 \) and \( a_2 \) are the coefficients of the linear regression between the air temperature and equilibrium temperature, \( K' = K/\rho \cdot C_w \) the modified exchange coefficient \( (\text{s}^{-1}) \) and \( B/A \) is approximated by \( 1/h \) where \( h \) is the water depth (m). Where water depths are not monitored, \( h \) can be estimated as a function of discharge: \( h = aQ^b \) (Leopold et al., 1964).

The equilibrium temperature values for the days where all needed meteorological data were available (from 1992 to 2005) were first calculated and the linear association between air and equilibrium temperatures was assessed. Once this association was verified, equation (5) was used to establish the model which will be further referred to as the EQB model.

### Stochastic autoregressive models with exogenous inputs

Stochastic autoregressive models consist of splitting water temperature series into two components that are then modeled adequately. For instance, water temperature may be divided
as follows:

\[ T_w(t) = TA_w(t) + R_w(t). \]  
(6)

The first component is the long-term annual component and represents the seasonal variations, and the second (residuals from the annual component) represents the short-term variations which are stationary. Using this approach a time series model can be fitted to water temperature residuals. Numerous time series models can be found in the literature such as Box-Jenkins, AR, ARMA, PAR, etc. (Benyahya et al., 2007a, 2007b).

The seasonal component is often modeled by a Fourier series analysis (Kothandaraman, 1971; El-Jabi et al., 1995) or even a simpler sinusoidal function (Cluis, 1972; Caissie et al., 1998). In this paper the Fourier series analysis was used to model both the air and water temperature seasonal components. Thus, these two functions are expressed as follows:

\[ TA(t) = \bar{T} + \sum_{k=1}^{\infty} \left( \chi_k \cos \left( \left( t - j_{T} + 1 \right) \frac{2\pi k}{N_{T}} + \phi_k \right) \right) \]  
(7)

where \( \bar{T} \) is the mean – water or air – temperature of the period \( T \), \( k \) is the order of each harmonic of the Fourier analysis, \( j_{T} \) is the rank of the first day where data is available in the period \( T \), \( N_{T} \) is the number of days of the period \( T \) and \( \chi_k \) and \( \phi_k \) are, respectively, the amplitude and the phase of each harmonic:

\[ \chi_{k^2} = A_{k^2} + B_{k^2}, \]
\[ \cos(\phi_k) = A_k/\chi_k \]

where \( A_k \) and \( B_k \) are Fourier coefficients that are expressed as follows:

\[ A_k = 2 \frac{N_{T}}{2 \pi} \sum_{t=1}^{N_{T}} \left( T(t) \cos \left( \frac{2\pi k t}{N_{T}} \right) \right) \]
\[ B_k = 2 \frac{N_{T}}{2 \pi} \sum_{t=1}^{N_{T}} \left( T(t) \sin \left( \frac{2\pi k t}{N_{T}} \right) \right). \]

Using the first harmonic \( (k = 1) \) to describe the long-term variations in air and water temperatures showed only small losses, and using the first two harmonics \( (k = (1, 2)) \) is sufficient, as reported by Kothandaraman (1971). Moreover, Kothandaraman also reported that up to 95% of the deviance can be explained by seasonal variations for water temperatures and up to 80% for air temperatures.

The residuals of the water temperatures were modeled by a second-order Markov process as suggested by Cluis (1972). The general form of the complete model is as follows:

\[ R_w(t) = A_1 R_w(t-1) + A_2 R_w(t-2) + K R_a(t) + \epsilon_1(t) \]  
(8)

where \( A_1 = R_1 \left( 1 - R_2 \right) / \left( 1 - R_1^2 \right) \) and \( A_2 = \left( R_2 - R_1^2 \right) / \left( 1 - R_1^2 \right) \) with \( R_1 \) and \( R_2 \) the autocorrelation coefficients for lags of 1 and 2 days. \( \epsilon_1 \) is the residual estimation error for this model and \( K \) represents the linear regression coefficient between the remaining residuals of the Markov process and the residuals of air temperature after removing the seasonal component, also referred to as the thermal exchange coefficient. This coefficient depends on many parameters such as stream cover, depth of water, etc. The value of this coefficient was estimated by the method of least squares. This model will be further referred to as the SMP1 model.

Equation (8) only takes into account the residual of air temperatures with no lag. However, Kothandaraman (1971) reported that residuals of air temperature with lags of up to two days were significant in explaining the evolution of water temperature residuals on a daily basis. Thus, the third model used in this study – which will be further referred to as SMPM – extends the Cluis approach by taking lagged air temperatures as predictors. The formulation of this model is therefore:

\[ R_w(t) = A_1 R_w(t-1) + A_2 R_w(t-2) + \sum_{i=1}^{\rho} K_i R_a(t-i) + \epsilon_2(t) \]  
(9)
where $\epsilon_2$ is the residual estimation error for this model and $K_i$ represents the linear regression coefficient between the residuals of the Markov process and the residuals of air temperatures with a lag of $i$ day(s). The maximum lag taken into account $p$ was first estimated using cross-correlation analysis between the residuals of the Markov process and the residuals of air temperature, and finally determined using the AIC criterion (Akaike, 1974). The method of least squares was used to estimate the values of the $K_i$.

The SMP1 and SMPM models only required water temperature and air temperature time series, which were available for the whole period 1978–2005. To conduct a fair comparison between these models and the equilibrium-based model, the same calibration and validation periods were used.

**Model calibration and performance assessment**

The data needed to establish all models were available for the years 1988 to 2005. This period was split into calibration and validation periods; respectively, 1988–1996 and 1997–2005. Data of the first part (1988–1996) were used to calibrate the models. Parameters were optimized using the least-squares method and using only data where observed temperatures were above 20°C to ensure better performance for high temperature prediction. The remaining years’ (1997–2005) data were afterwards used for validation of the models.

To assess the performance of each model in predicting the daily mean water temperatures, the root mean square error criterion was used (RMSE (Berger, 1985)) that is calculated by:

$$
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (T_{OBS,i} - T_{PRE,i})^2}
$$

(10)

where $N$ is the number of observations, $T_{OBS,i}$ the observed daily mean water temperatures and $T_{PRE,i}$ the predicted daily mean water temperatures.

Another criterion was used during the validation step that is more related to the conditions of salmon migration and viability. Three important temperature thresholds were selected that reflect migrating conditions: 9°C, 19°C and 24°C. Chanseau et al. (1999) showed that 9°C and 24°C are, respectively, the lower and upper limits for salmon migration, with no passage at fish passage facilities for lower or higher temperatures. Above 24°C, salmons are sensitive to thermal stress and most of the salmon mortalities in the Garonne River were recorded for such temperatures (Croze et al., 2006). The last temperature threshold (19°C) was reported as the upper limit for optimal conditions for youth growth (Decola, 1970; Swansburg et al., 2002). The performance of each model regarding these thresholds was assessed by comparing the temporal occurrence and duration of periods between consecutive thresholds of observed and predicted time series.

**DATA AND STUDY AREA**

> **STUDY AREA**

The study area is located on the Garonne River at the Malause reservoir, upstream of the headrace of the Golfech nuclear power plant (Figure 1). The water thermal regime in this location is influenced by climatic conditions, and the Tarn tributary which has a junction with the Garonne River is located 5 km upstream of the study site. The climate is temperate and characterized by mostly gentle winters and hot summers. The Garonne between Toulouse and the study site is a well-mixed, wide and shallow river; therefore profiles of water temperatures were assumed to be laterally and vertically uniform.
Data

Along our study reach a large amount of climatic and hydrologic data was gathered. Daily mean water temperature series in Malause were available from 1978 to 2005 from measurements in the headrace of the Golfech nuclear power plant performed by EDF\(^1\). Hydrological data were available at the Lamagistère station of the Banque HYDRO\(^2\) that is 17 km downstream of Malause. Mean daily discharges were available from 1967 with no uncertain or missing data and mean daily water levels from 1988 with about 1% of uncertain data and less than 0.01% of missing data. Three-hourly meteorological data were gathered at the Agen and Blagnac weather stations of Météo-France\(^3\). Air temperatures were available at both stations for the years 1978 to 2005 with less than 0.01% of missing or uncertain data. Incident solar radiation (also referred to as insolation) data were only available at the Agen station, with few missing data (0.2%) from 1978 to 2005. Cloud cover and wind speed were only available at the Blagnac station from 1992 but with numerous uncertain or missing data. Cloud cover uncertain data were about 8% (no missing data) and the wind speed time series presents

---


\(^2\) Banque HYDRO, [http://www.hydro.eaufrance.fr](http://www.hydro.eaufrance.fr).

\(^3\) Météo-France, [http://www.meteofrance.com](http://www.meteofrance.com).
Table I
Availability and consistency of daily mean data.

<table>
<thead>
<tr>
<th>Station</th>
<th>Availability</th>
<th>Validate</th>
<th>Doubtful</th>
<th>Missing</th>
<th>Producer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discharge</td>
<td>Lamagistère</td>
<td>1967–2005</td>
<td>100%</td>
<td>0%</td>
<td>Banque HYDRO</td>
</tr>
<tr>
<td>Water level</td>
<td>Lamagistère</td>
<td>1988–2005</td>
<td>100%</td>
<td>0%</td>
<td>Banque HYDRO</td>
</tr>
<tr>
<td>Water temperature</td>
<td>Malaguse</td>
<td>1978–2005</td>
<td>100%</td>
<td>0%</td>
<td>EDF</td>
</tr>
<tr>
<td>Air temperature</td>
<td>Blagnac</td>
<td>1978–2006</td>
<td>100%</td>
<td>0%</td>
<td>Météo-France</td>
</tr>
<tr>
<td>Air temperature</td>
<td>Agen</td>
<td>1978–2006</td>
<td>100%</td>
<td>0%</td>
<td>Météo-France</td>
</tr>
<tr>
<td>Incoming solar radiation</td>
<td>Agen</td>
<td>1992–2005</td>
<td>100%</td>
<td>0%</td>
<td>Météo-France</td>
</tr>
<tr>
<td>Nebulosity</td>
<td>Blagnac</td>
<td>1992–2005</td>
<td>100%</td>
<td>0%</td>
<td>Météo-France</td>
</tr>
<tr>
<td>Wind speed</td>
<td>Blagnac</td>
<td>1992–2005</td>
<td>96%</td>
<td>4%</td>
<td>Météo-France</td>
</tr>
</tbody>
</table>

Table II
Analysis of trends and significance (p-value) in water temperatures and climatic parameters with significant trends.

<table>
<thead>
<tr>
<th></th>
<th>Water temperature</th>
<th>Air temperature</th>
<th>Incident solar radiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evolution</td>
<td>p-Value</td>
<td>Evolution</td>
<td>p-Value</td>
</tr>
<tr>
<td>5th percentile</td>
<td>–</td>
<td>0.493</td>
<td>–</td>
</tr>
<tr>
<td>95th percentile</td>
<td>+2.12 °C</td>
<td>0.010</td>
<td>+2.79 °C</td>
</tr>
<tr>
<td>Annual average</td>
<td>+1.29 °C</td>
<td>0.007</td>
<td>+1.58 °C</td>
</tr>
<tr>
<td>Winter average</td>
<td>–</td>
<td>0.338</td>
<td>–</td>
</tr>
<tr>
<td>Spring average</td>
<td>+1.60 °C</td>
<td>0.003</td>
<td>+2.15 °C</td>
</tr>
<tr>
<td>Summer average</td>
<td>+2.92 °C</td>
<td>0.001</td>
<td>+2.57 °C</td>
</tr>
<tr>
<td>Fall average</td>
<td>–</td>
<td>0.446</td>
<td>–</td>
</tr>
</tbody>
</table>

RESULTS

> TRENDS AND CORRELATION

Trend analyses of seven descriptive statistics of water temperatures, and hydraulic and climatic parameter time series were performed: 5th and 95th annual percentiles, and annual and seasonal averages. Cloud cover, wind speed and hydraulic parameters showed no significant trend at all. The results for the remaining parameters are listed in Table II and respective evolutions are plotted in Figure 2. Significant trends at the 99% confidence interval (p < 0.01) were found for the 95th percentile and for the annual, spring and summer averages of both air and water temperatures. The annual averages of water temperatures in Malaguse have risen by 1.29 °C and summer averages by 2.92 °C. Air temperatures showed similar trends, with annual averages that have risen by 1.58 °C and summer averages by 2.57 °C. Significant trends for insolation were only found for the annual and spring averages, with rises of 28 J·cm⁻²·day⁻¹ and 55 J·cm⁻²·day⁻¹, respectively.

Correlation analyses on water temperatures against air temperatures and incident solar radiation were performed afterwards for the four descriptive statistics with significant trends. Analyses were performed using one parameter at a time (Table III). At the 99% confidence
Figure 2

Trends in water temperatures in Malause (Figure 2.a), air temperatures (Figure 2.b) and incident solar radiation (Figure 2.c) at the Agen station.

Table III

Analysis of correlations and significance of water temperatures against air temperatures and incident solar radiation.

<table>
<thead>
<tr>
<th></th>
<th>Air temperature</th>
<th>Incident solar radiation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation</td>
<td>95th percentile</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>Annual average</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>Spring average</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Summer average</td>
<td>0.87</td>
</tr>
</tbody>
</table>

interval, air temperature was the most significant predictor, that explained more than 75% of the variance of the water temperature descriptive statistics. Regarding high water temperatures (summer averages or the 95th percentile), air temperatures explained more than 85% of the variance. The lowest calculated correlation coefficient was for spring averages ($R^2 = 0.78$). Insolation was significantly correlated with water temperatures only for spring averages and summer averages (at the 99% confidence interval) but with correlation coefficients of less than 0.60. The best correlation coefficient was obtained for spring averages ($R^2 = 0.58$, $p = 0.001$).
Table IV
RMSE (°C) calculated from observed and predicted daily mean water temperatures using the EQB model for the years 1988–2005 in Malause.

Tableau IV
Erreurs-types (°C) calculées entre les observations des températures moyennes journalières de l’eau et les estimations fournies par le modèle EQB sur la période 1988–2005.

<table>
<thead>
<tr>
<th>Year</th>
<th>$T_{OBS}$ &gt; 20 °C</th>
<th>Whole year $T_{OBS}$ &gt; 20 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>1.43</td>
<td>1.23</td>
</tr>
<tr>
<td>1989</td>
<td>1.17</td>
<td>1.07</td>
</tr>
<tr>
<td>1990</td>
<td>0.89</td>
<td>0.76</td>
</tr>
<tr>
<td>1991</td>
<td>1.19</td>
<td>1.25</td>
</tr>
<tr>
<td>1992</td>
<td>1.54</td>
<td>1.05</td>
</tr>
<tr>
<td>1993</td>
<td>1.35</td>
<td>1.40</td>
</tr>
<tr>
<td>1994</td>
<td>1.10</td>
<td>1.56</td>
</tr>
<tr>
<td>1995</td>
<td>0.89</td>
<td>1.41</td>
</tr>
<tr>
<td>1996</td>
<td>1.26</td>
<td>1.61</td>
</tr>
<tr>
<td>1988–1996</td>
<td>1.22</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Calibration period Validation period

> EQUILIBRIUM TEMPERATURE-BASED MODEL

Mean daily equilibrium temperatures showed a strong linear relation with daily mean air temperatures in Malause, with a calculated value of $R^2 = 0.99$. The EQB model was therefore calibrated using data from the calibration period (1988–1996). The coefficients of the linear regression between estimated daily mean equilibrium temperatures and daily mean air temperatures were estimated as $a_1 = 1.12$ and $a_2 = 0.44$ °C. Using these values, the modified thermal coefficient was calculated at $K' = 0.56 \text{ s}^{-1}$.

On both the calibration period and validation period, RMSE were calculated for the overall period and for each year, both using all data and data with high observed daily mean water temperatures (above 20 °C) only (Table IV). On the calibration period, the overall RMSE was calculated at 1.22 °C and 0.81 °C for temperatures above 20 °C. Inter-annual comparison showed that the RMSE ranged from 0.89 °C (1990, 1995) to 1.54 °C (1992) using all data and 0.53 °C (1995) to 1.15 °C (1993) for temperatures above 20 °C. On the validation period, the RMSE was calculated at 1.31 °C. Values calculated for each year ranged from 1.05 °C (2001) to 1.61 °C (2005). Regarding temperatures above 20 °C only, the RMSE ranged from 0.61 °C (1999) to 1.72 °C (2005) with an overall value calculated at 1.22 °C. The fitness of the results can also be assessed in Figure 4. The model showed good agreement on temperatures above 20 °C, except for the year 2005 where temperatures around day 180 were overpredicted. Overpredictions were also noted for temperatures just below 20 °C, as for spring of the year 1998.

> STOCHASTIC MODELS

The values calculated for $x_k$ and $\phi_k$ for the first two harmonics of equation (7) are listed in Table V. Comparison of the seasonal components of water and air temperatures indicated that water temperatures – apart from short-term variations – are always higher than air temperatures (see Figure 3). Moreover, the relation between interannual means was $T_w = 1.12 T_a + 0.37$. The seasonal component of water temperatures explained 92% of the deviance in the period 1988–1996 and 84% in the period 1997–2005. Regarding air temperatures, the proportion of deviance explained by the seasonal components was 76% in the period 1988–1996 and 70% in the period 1997–2005. These values agree with those reported by Kothandaraman (1971). The maximum temperature for the long-term component of the stream water temperature was reached on day 226 (August 14) at a value of 23.8 °C.
Once the long-term variations were removed from the water temperature time series, the SMP1 and SMMP models were fitted on the residuals. The autocorrelation coefficients for the water temperature residuals were calculated at $r_1 = 0.97$ and $r_2 = 0.91$. The Markov coefficients were therefore calculated at $A_1 = 1.54$ and $A_2 = -0.58$.

The thermal exchange coefficient of the SMP1 model was then calculated as 0.049 using data for the validation period (1988–1996). The SMP1 model is therefore expressed as follows:

$$R_w(t) = 1.54R_w(t - 1) - 0.58R_w(t - 2) + 0.049R_a(t) + \varepsilon_1(t). \tag{11}$$

Root mean square errors between predicted and observed values were calculated on the calibration period as well as for each year and for high observed water temperatures (Tables VI and VII). The RMSE was calculated as 1.13 °C when using all data and 0.95 °C for high temperatures. Overall results were better than those obtained with the EQB model but RMSE calculated for high temperatures were a little bit higher. Results also varied from year to year, with the RMSE ranging from 0.87 °C (1994) to 1.35 °C (1988) for overall data and from 0.49 °C (1995) to 1.32 °C (1993) for high temperatures. On the validation period, the RMSE ranged from 1.11 °C (1999, 2001) to 1.49 °C (2005) using all data and from 0.72 °C (1999) to 1.72 °C (2003) for temperatures above 20 °C, with values of 1.27 °C calculated on the whole validation period using both all data and temperatures above 20 °C.

To establish the SMMP model, cross-correlation analysis between the residuals of the Markov process and the residuals of air temperatures revealed an exponential decrease in the cross-correlation coefficient when increasing the lag. The cross-correlation coefficients calculated for lags of 0 to 3 days were, respectively, 0.53, 0.41, 0.32 and 0.25. The analyses of the AIC criterion and significance of the regression coefficients showed that lags higher than 3 days were insignificant. Finally, the SMMP model is expressed as follows:

$$R_w(t) = 1.54R_w(t - 1) - 0.58R_w(t - 2) + 0.045R_a(t)$$
$$+0.035R_a(t - 1) - 0.032R_a(t - 2) - 0.014R_a(t - 3) + \varepsilon_2(t). \tag{12}$$

This model clearly provided better results than the SMP1 model (Table VII). Moreover, it fitted better than the EQB model on the whole time series with the RMSE calculated at 1.04 °C.
Table V
Coefficient values of seasonal (Eq. (10)) components for air and water temperatures.

Tableau V
Coefficients des composantes saisonnières (Éq. (7)) des chroniques de températures de l’air et de l’eau.

<table>
<thead>
<tr>
<th></th>
<th>Water temperature</th>
<th>Air temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi_1$</td>
<td>8.43</td>
<td>7.93</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>2.67</td>
<td>2.79</td>
</tr>
<tr>
<td>$\chi_2$</td>
<td>1.48</td>
<td>0.84</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-1.73</td>
<td>-1.50</td>
</tr>
</tbody>
</table>

Table VI
RMSE (°C) calculated from observed and predicted daily mean water temperatures using the SMP1 model and the SMPM model for the years 1988–2005 in Malause.

Tableau VI
Erreurs-types (°C) calculées entre les observations des températures moyennes journalières de l’eau et les estimations fournies par les modèle SMP1 et SMPM sur la période 1988–2005.

<table>
<thead>
<tr>
<th>Year</th>
<th>SMP1</th>
<th>SMPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>1.35</td>
<td>1.34</td>
</tr>
<tr>
<td>1989</td>
<td>1.03</td>
<td>0.87</td>
</tr>
<tr>
<td>1990</td>
<td>1.04</td>
<td>0.89</td>
</tr>
<tr>
<td>1991</td>
<td>1.11</td>
<td>1.03</td>
</tr>
<tr>
<td>1992</td>
<td>1.22</td>
<td>1.45</td>
</tr>
<tr>
<td>1993</td>
<td>1.27</td>
<td>0.83</td>
</tr>
<tr>
<td>1994</td>
<td>0.87</td>
<td>0.80</td>
</tr>
<tr>
<td>1995</td>
<td>0.95</td>
<td>0.90</td>
</tr>
<tr>
<td>1996</td>
<td>1.20</td>
<td>1.06</td>
</tr>
<tr>
<td>1988–1996</td>
<td>1.13</td>
<td>1.04</td>
</tr>
<tr>
<td>1997</td>
<td>1.34</td>
<td>1.14</td>
</tr>
<tr>
<td>1998</td>
<td>1.15</td>
<td>0.92</td>
</tr>
<tr>
<td>1999</td>
<td>1.11</td>
<td>1.18</td>
</tr>
<tr>
<td>2000</td>
<td>1.13</td>
<td>1.09</td>
</tr>
<tr>
<td>2001</td>
<td>1.11</td>
<td>0.86</td>
</tr>
<tr>
<td>2002</td>
<td>1.32</td>
<td>1.44</td>
</tr>
<tr>
<td>2003</td>
<td>1.44</td>
<td>1.17</td>
</tr>
<tr>
<td>2004</td>
<td>1.29</td>
<td>1.39</td>
</tr>
<tr>
<td>2005</td>
<td>1.49</td>
<td>1.43</td>
</tr>
<tr>
<td>1997–2005</td>
<td>1.27</td>
<td>1.20</td>
</tr>
</tbody>
</table>

on the calibration period (versus 1.20 °C for the EQB model and 1.13 °C for the SMP1 model) and 1.20 °C on the validation period (versus 1.31 °C for the EQB model and 1.27 °C for the SMP1 model). Results for temperatures above 20 °C were about the same as those obtained with the EQB model on the calibration period (0.79 °C versus 0.81 °C for the EQB model) and were better than with the SMP1 model (0.95 °C). The RMSE ranged for this period between 0.80 °C (1994) and 1.45 °C (1992) using all data and from 0.55 °C (1995) to 1.02 °C (1992) using temperatures above 20 °C only. For the validation period the SMPM model clearly provided better results for temperatures above 20 °C, with the RMSE calculated at 1.01 °C (versus 1.22 °C for the EQB model and 1.27 °C for the SMP1 model) and ranging from 0.56 °C (1999) to 1.40 °C (2002). Comparison between the EQB model and SMPM model predictions for the years 1992 and 2005 (Figures 4 and 5) clearly showed the better accuracy of the SMPM model at low temperatures. The end of the 2005 time series was particularly better predicted by the SMPM model than by the EQB model.
Table VII
RMSE (°C) calculated from observed and predicted daily mean water temperatures using only the seasonal component, the SMP1 model and the SMPM model for the years 1988–2005 in Malause.

<table>
<thead>
<tr>
<th>Year</th>
<th>Seasonal component</th>
<th>SMP1</th>
<th>SMPM</th>
<th>1997</th>
<th>Seasonal component</th>
<th>SMP1</th>
<th>SMPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>1.53</td>
<td>1.35</td>
<td>1.34</td>
<td>2.05</td>
<td>1.34</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td>1.60</td>
<td>1.03</td>
<td>0.87</td>
<td>1.55</td>
<td>1.15</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>2.01</td>
<td>1.04</td>
<td>0.89</td>
<td>1.37</td>
<td>1.11</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>1.87</td>
<td>1.11</td>
<td>1.03</td>
<td>2.00</td>
<td>1.52</td>
<td>1.13</td>
<td>1.09</td>
</tr>
<tr>
<td>1992</td>
<td>2.31</td>
<td>1.22</td>
<td>1.45</td>
<td>2.01</td>
<td>1.90</td>
<td>1.11</td>
<td>0.86</td>
</tr>
<tr>
<td>1993</td>
<td>1.66</td>
<td>1.27</td>
<td>0.83</td>
<td>2.02</td>
<td>1.72</td>
<td>1.32</td>
<td>1.44</td>
</tr>
<tr>
<td>1994</td>
<td>1.61</td>
<td>0.87</td>
<td>0.80</td>
<td>2.03</td>
<td>2.04</td>
<td>1.44</td>
<td>1.17</td>
</tr>
<tr>
<td>1995</td>
<td>1.62</td>
<td>0.95</td>
<td>0.90</td>
<td>2.04</td>
<td>1.54</td>
<td>1.29</td>
<td>1.39</td>
</tr>
<tr>
<td>1996</td>
<td>1.69</td>
<td>1.20</td>
<td>1.06</td>
<td>2005</td>
<td>2.19</td>
<td>1.49</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Validation period | Calibration period

Figure 4
Results for modeling daily mean water temperatures by the EQB model in Malause for both the calibration period (Figures 4.a and 4.b) and validation period (Figures 4.c and 4.d).

Figure 4
Estimations des températures de l’eau à Malause par le modèle EQB, sur la période de calibration (Figures 4.a et 4.b) et sur la période de validation (Figures 4.c et 4.d).
Figure 5

Results for modeling daily mean water temperatures by the stochastic SMPM model in Malause for both the calibration period (Figures 5.a and 5.b) and validation period (Figures 5.c and 5.d).

> TEMPERATURE CONDITIONS FOR MIGRATION

The last evaluation of those models consisted of evaluating their propensity to predict the crossing of the three thresholds of water temperature related to salmon conditions for migration and viability. The time locations of the corresponding periods are plotted in Figure 6. In this figure, short-lasting threshold crossings (less than 7 days) were erased to improve clarity. Both models showed good accuracy, particularly in predicting the 24 °C threshold crossings. Some large differences were, however, noted, as for the first crossing of the 19 °C threshold in the year 1999. Distribution of errors between predicted and observed temperature were calculated for each threshold and for days where crossings were not well predicted (Figure 7). Except for the 19 °C threshold, about 80% of errors were in the range [-2 °C; 2 °C]. For the 9 °C and 24 °C thresholds, SMPM performed slightly better with, respectively, 87% and 84% of absolute errors less than 2 °C (versus 82% and 77% for the EQB model). On the contrary, the 9 °C threshold crossing was better predicted by the EQB with 57% of absolute errors less than 2 °C (versus 52% for the SMPM model).

DISCUSSION

Water temperatures in streams can be related to numerous factors such as climate, hydraulic regimes, bed topography, and others (Caissie, 2006; Webb et al., 2008). Deterministic models using many factors have been used in the literature and proved to efficiently predict water temperatures (Sinokrot and Stefan, 1984; Kim and Chapra, 1997; Webb and Zhang, 1999;
Marcé and Armengol, 2008). Using those models, however, requires lots of data and computational resources. Such models are consequently not easy to use for practitioners. This study was therefore conducted on the Garonne River to assess the performance of statistical models to predict daily mean water temperatures and particularly high temperatures that impact on aquatic ecosystems.

Trend analyses revealed that water temperature evolution was closely similar to that of air temperatures. Similar results were reported for two large rivers in France, the Loire River (Moatar and Gailhard, 2006) and the Rhône River (Poirel et al., 2008). Such similarities tend to indicate that water temperature in large rivers in France is mainly influenced by climatic conditions and particularly air temperatures. Solar radiation was also noted to be correlated with the water temperature but using this factor as a predictor would potentially have resulted in statistical inadequacies associated with multicollinearity. Therefore, using models relying on the relation between air and water temperatures seemed to be accurate for the Garonne River case.

The first model used in this study was based on the equilibrium temperature concept. The equilibrium temperature reflects the energy budget of the stream and therefore is a function of many meteorological factors. It has been shown that the equilibrium temperature could be expressed as a simple linear function of air temperatures for temperate regions (Caissie et al., 2005). This assumption was verified for the Garonne River, with good agreement between air temperatures and the equilibrium temperature ($R^2 = 0.99$). The $a_1$ coefficient was optimized at a value of 1.12. This value was similar to that reported by Caissie et al. (2005) for the Little
Southwest Miramichi River (New Brunswick, Canada). Values of this coefficient higher than 1 reflect that the river is well exposed to other meteorological factors than air temperatures (i.e. solar radiation, etc.) which was the case of the Garonne river due to its wideness. The $a_2$ coefficient, however, was not zero, which differs from the results reported by Caissie et al. The thermal coefficient $K'$ was calculated at 0.71 and agreed with that of the Little Southwest Miramichi River.

The results obtained with the equilibrium-based model were similar to those reported in other studies, with values of RMSE calculated at 1.22 °C on the calibration period and 1.31 °C on the validation period. Regarding temperatures above 20 °C only, slightly better values of 0.81 °C on the calibration period and 1.22 °C on the validation period were obtained. This model was therefore sensitive to the data used for calibration. Inter-annual comparison also indicated that estimations for several years were poorly estimated. Except for the year 1988, whose measured water temperature time series contained outliers (caused by measurement failures), correlations between the model residuals and water depths were found to be strong for those years, whereas poor correlations were calculated for years with good results (Table VIII). The EQB model was therefore sensitive to water depths and approximating $B/A$ with $1/h$ could be debated. Another simplification – neglecting diffusive and convective terms – could also be too restrictive. Significant floods with discharge of more than 5000 m$^3$·s$^{-1}$ happen in this river. During such floods, the importance of convective terms must not be negligible.

Stochastic modeling consisted of separating water temperatures and air temperatures into two components: a seasonal component (long-term variations) and residuals (short-term
Table VIII
Correlation between the EQB model residuals and water depths.

<table>
<thead>
<tr>
<th>Year</th>
<th>Correlation</th>
<th>p-value</th>
<th>Year</th>
<th>Correlation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>-0.19</td>
<td>0.000</td>
<td>1997</td>
<td>-0.11</td>
<td>0.044</td>
</tr>
<tr>
<td>1989</td>
<td>-0.21</td>
<td>0.000</td>
<td>1998</td>
<td>-0.17</td>
<td>0.001</td>
</tr>
<tr>
<td>1990</td>
<td>-0.15</td>
<td>0.389</td>
<td>1999</td>
<td>-0.24</td>
<td>0.000</td>
</tr>
<tr>
<td>1991</td>
<td>-0.69</td>
<td>0.000</td>
<td>2000</td>
<td>-0.16</td>
<td>0.003</td>
</tr>
<tr>
<td>1992</td>
<td>-0.57</td>
<td>0.000</td>
<td>2001</td>
<td>-0.19</td>
<td>0.000</td>
</tr>
<tr>
<td>1993</td>
<td>-0.57</td>
<td>0.083</td>
<td>2002</td>
<td>-0.36</td>
<td>0.000</td>
</tr>
<tr>
<td>1994</td>
<td>-0.16</td>
<td>0.037</td>
<td>2003</td>
<td>-0.44</td>
<td>0.000</td>
</tr>
<tr>
<td>1995</td>
<td>-0.11</td>
<td>0.166</td>
<td>2004</td>
<td>-0.56</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Validation period | Calibration period

Second-order Fourier series were fitted to inter-annual daily means of water and air temperatures to establish the seasonal components. Values obtained for coefficients of the Fourier series used to model the water and air temperature seasonal component indicated that seasonal amplitude of water temperatures was always higher than that of air temperatures – which reflected the influence of other meteorological factors as reported for the EQB model – and that air temperatures have on average a 7-day delay from water temperatures, which was due to the thermal inertia of the water.

Once the seasonal component was removed from water temperatures, the residuals of water temperature were modeled using two stochastic models. These two models used a second-order autoregressive model with exogenous variables. The SMP1 only used air temperature residuals as an exogenous predictor, while the SMPM model used lagged (up to 3 days) air temperature residuals. Coefficients of the autoregressive model were calculated using autocorrelation coefficients of the water temperature residuals. Autocorrelation coefficients were similar to those reported for large rivers, for instance $R_1 = 0.84$ in Illinois river, IL, USA (Kothandaraman, 1971) and $R_1 = 0.92$ in the Rivière du Nord near Montreal, P.Q., Canada (Cluis, 1972). Finally, the remaining coefficients of both models were optimized using measured water temperatures above $20^\circ$C.

The stochastic models were more robust than the EQB model due to the fitness of the seasonal curve. As more than 80% of the variance of water temperature was explained by the seasonal component, the stochastic models’ performance mainly depended on the fitness of this component. The SMP1 model performed slightly better than the EQB model when comparing RMSE using all data. On the contrary, RMSE calculated for temperatures above $20^\circ$C were slightly higher. The results obtained for the SMPM model were on average better than those obtained with the other two models. As for the EQB model, water temperatures for several years were poorly predicted, such as the years 1988 and 1992 in the calibration period and the years 2002, 2004 and 2005 in the validation period. Except for the year 1988 (outliers) the seasonal component for these years were poorly fitted (Table VII), which explained the poor results obtained by the stochastic models.

The performance of the EQB and SMPM models was also tested on predicting the crossing of water temperature thresholds linked to the conditions of Atlantic Salmon. Despite several differences, the models showed good performance. Analyses of errors for days where threshold crossings were not well predicted revealed that most of these errors were in the range $[-2^\circ$C; $2^\circ$C]. Errors for the $19^\circ$C crossing were, however, bigger; especially for the SMPM model, that performed more poorly for temperatures just under $20^\circ$C.
CONCLUSION

Both approaches used in this study could be useful for practitioners. Despite being applied to a complex study area (reservoir and influence of tributary), these simplified models showed good accuracy in predicting high water temperatures of the Garonne River in Malause. As water and air temperatures are relatively inexpensive to measure, statistical models are good alternatives to deterministic models that require much more data. In our study, equilibrium temperatures were established from meteorological parameters, and values of the linear regression between equilibrium temperatures and air temperatures were calculated afterwards. However, it should be possible to establish the model directly using only air and water temperatures and an optimization method in order to use only air temperatures, water temperatures and water depths. Furthermore, using real values of top width and wetted area or non-linear regression would probably result in better prediction, as well as multiplying each thermal flux by a calibration factor (Caissie et al., 2007) to slightly modify their respective influences. Regarding the stochastic models, performance was mainly dependent on the fitness of the seasonal component. Therefore, further research is needed to explore the advantages of using variable coefficients for this component. Finally, as the meteorological station was located 17 km downstream of the study site, potential improvements could also be made by alternatively using microclimate data (Benyahya et al., 2010), particularly in determination of heat fluxes for the EQB model.

Each model could have different usage for practitioners. As the EQB predicts water temperatures from climatic conditions, this model could be useful to assess the evolution of the thermal regime of the Garonne River under climate change. Using regional data derived from Global Circulation Model (GCM) outputs, water temperatures could be calculated for future years. On the contrary, the stochastic models used in this study require knowledge of past water temperature values. These models are therefore more suitable for short-term predictions.

REFERENCES


