دومین کنفرانس سراسری آبخیزداری و مدیریت منابع آب و خاک
کرمان ۳ و ۴ اسفندماه ۱۳۸۴
دانشگاه شهید باهنر
انجمن مهندسی آبیاری و آب ایران

ارزیابی کارایی هوش مصنوعی کامپیوتر در تخمین داده های مفقود شده هیدرولوژی

Evaluation of the applicability of artificial intelligence on hydrological missing data estimation

محمدتقی دستورانی[۱]، استادیار دانشگاه یزد

چکیده :

در این پروژه تحقیقاتی سعی شده است که مقادیر مربوط به نواقص آماری ایستگاهها با استفاده از مقادیر ایستگاههای دیگر موجود در گروه مربوطه و به کمک ساختار مناسبی از شبکه های عصبی مصنوعی برآورد گردد . جهت امکان مقایسه این روش با روشهای موجود نتایج حاصل از شبکه های عصبی مصنوعی با نتایج روش نسبت نرمال و روش همبستگی بین ایستگاهها مورد بررسی و مقایسه قرار گرفته است. بدین منظور مقادیر مربوط به ریشه میانگین مربع خطا (***RMSE***) و ضریب کارایی(R^2) برای نتایج ارائه شده توسط هریک از روشها و مقادیر واقعی (اندازه گیری شده) محاسبه گردید. بر اساس نتایج به دست آمده از این تحقیق میتوان گفت که تکنیک شبکه های عصبی در حقیقت کارایی مناسب جهت پیش بینی داده های مفقود شده هیدرولوژی را دارد. در بیشتر موارد در این تحقیق درستی نتایج حاصل از شبکه های عصبی مصنوعی بالاتر از نتایج بدست آمده از دو روش دیگر بوده است . شبکه عصبی مصنوعی از نوع پرسپترون چند لایه سازگار ترین ساختار برای این هدف بوده است هرچند که در موارد معدودی شبکه عصبی برگشتی نیز نتایجی همانند شبکه پرسپترون چند لایه ویا حتی اندکی بهتر از آنرا ارائه نموده است .

Abstract:

In this research project, it was tried to predict the missing data of each station using data from other stations of the same group, and a relevant architecture of the artificial neural networks model. To be able to compare the results produced by artificial neural network models, the normal ratio method and the correlation method were also employed to the same job (to predict missing data predicted by ANN technique in the first stage). Then the values of root mean square error (rmse) and coefficient of efficiency (r^2) was determined for the predicted values of each method and the corresponding measured values. According to the results taken from this research it can be said that artificial

[۱]- یزد، صفاییه، دانشگاه یزد، دانشکده منابع طبیعی و کویرشناسی
تلفاکس:۰۳۵۱-۸۲۱۰۳۱۲ E-mail:mdastorani@yazduni.ac.ir

neural network is an efficient method to predict the missing data. In most of the cases in this research accuracy of the results produced by ANN method is considerably higher than those produced by other two approaches. Multi-layer perceptron was the most relevant architecture for this application, although in few cases the out puts of the recurrent network was similar or slightly better than the results presented by MLP (multi-layer perceptron) network.

۱-مقدمه

انجام هر نوع برنامه ریزی که از اطمینان و دقت کافی برخوردار باشد نیاز به آمار و اطلاعات اندازه گیری شده جریان آب در بسترهای طبیعی و رودخانه ها دارد چرا که بدون این آمار و اطلاعات برآورد میزان آب قابل دسترس ممکن نبوده و هر نوع برنامه ریزی در زمینه منابع آبی کشور اعتماد و دقت کافی را نخواهد داشت . این آمار و اطلاعات معمولاً در قالب داده های هیدرولوژیکی طبقه بندی و پردازش می شود که البته اغلب دارای نواقصی نیز می باشد که نیاز به بازسازی دارد . روشهای بازسازی آمار هیدرولوژیکی که به صورت کلاسیک موجود است هیچگاه بدون خطا نمی باشد که البته روشی که خطای کمتری داشته باشد کارایی بیشتری برای این هدف دارد .

در دهه های اخیر تکنیک های کامپیوتری جدید و از جمله هوش مصنوعی کامپیوتر که در قالب واریته های متفاوت و از جمله شبکه های عصبی مصنوعی *(Artificial Neural Networking)* جهت حل مسائل مختلف مورد استفاد ه قرار گرفته و اساسا در زمینه های مختلف تحقیقاتی نتایج درخشانی را نیز از خود نشان داده است. شبکه عصبی مصنوعی در واقع مدل ساده شده ای از مغز انسان می باشد. بدون شک مغز انسان پیشرفته ترین پردازشگری است که در طبیعت وجود دارد . عدم نیاز به الگوریتم و معادلات حاکم بر پدیده، امکان یادگیری از طریق ارائه مثال، قدرت تعمیم دهی بالا و ثبت اطلاعات در ساختاری کاملاً گسترده و به شکلی توزیع یافته از ویژگیهای منحصر به فرد آن است . شبکه عصبی مصنوعی در واقع یک ساختار ریاضی است که توانایی نشان دادن فرایندها و ترکیب دلخواه غیر خطی جهت ارتباط بین ورودیها و خروجیهای هر سیستم را دارا است . این شبکه با داده های موجود طی فرایند یادگیری آموزش دیده و جهت پیش بینی در آینده مورد استفاده قرار می گیرد. در این تحقیق کارایی این تکنیک به منظور بازسازی داده های هیدرولوژیکی مورد ارزیابی قرار گرفته والبته نتایج حاصل با روشهای موجود مقایسه گردیده است .

۲-نواقص آماری

یکی ار کاربردهای آمار در هیدرولوژی آن است که بتوانیم برخی از خصوصیات آب و هوایی یا هیدرولوژیکی مناطقی را که دارای داده های آماری کم یا اصولاً فاقد آمار هستند را تخمین بزنیم . زیرا در بسیاری موارد نمی توان اجرای یک پروژه را فقط به دلیل آنکه در مورد آن داده های هیدرولوژیکی دراز مدت وحود ندارد به تعویق انداخت . از

طرف دیگر نمی توان نقش داده ها را در هیدرولوژی نادیده گرفت . پس لازم است به روشهای مختلف داده های مورد نیاز را تخمین زد .

روشهای تخمین داده های هیدرولوژیکی در مورد تکمیل سریهای آماری نیز به کار برده می شود. موارد زیادی مشاهده می شود که بنا به دلایلی آمار یک روز یا یک ماه و سال خاص مفقود شده یا اصلاً برداشت نشده است ، تاریخ تأسیس ایستگاهها در یک منطقه با یکدیگر متفاوت می باشد و از طرف دیگر نقایص احتمالی دستگاهها و برداشت آمار غلط که توسط کارشناسان کنترل شده است واز مجموعه آمار حذف می شوند، ویا از بین رفتن ایستگاهها در اثر سوانح طبیعی مانند سیل ، زلزله و یا انهدام آنها در اثر جنگ و ... باعث کمبود و نقص در سریهای آماری شده و لذا لازم است این داده ها قبل از آنکه مورد تجزیه و تحلیل قرار بگیرند بایستی تکمیل شوند . روشهایی که برای این منظور به کار می روند متعدد است که از مهمترین آنها میتوان به روش نسبت نرمال و روش همبستگی بین ایستگاهها اشاره کرد.

به لحاظ برخی مشکلات در جمع آوری داده ها و بازسازی آنها احتیاج به ابزاری است که قابلیت درک و یادگیری توابع حاکم بر داده را داشته باشد و بعلاوه بتواند درمقابل نوسانهای ناشی از عدم دقت در داده مقاومت نماید . با توجه به پیچیدگی و طبیعت نوسانی داده های هیدرولوژیکی چنین به نظر می رسد که با در نظر گرفتن ویژگیها و قابلیتهای منحصر به فرد شبکه های عصبی مصنوعی این شبکه ها بتوانند به عنوان ابزاری مفید و کارآمد مورد استفاده قرار گیرند.

۳-شبکه عصبی مصنوعی

شبکه عصبی مصنوعی در حقیقت یک سیستم محاسباتی است که بر اساس سیستم عصبی جانداران (انسان) ساخته شده است. ساختمان این شبکه متشکل از نرونهای مصنوعی (همانند سلولهای عصبی جانداران) است که انتقال سیگنالها را که در اثر ورودیهای سیستم ایجاد شده انجام میدهد. این عمل همانند انتقال پیام در شبکه عصبی جانداران است که در اثر محرک خارجی ایجاد شده و با ترشح هورمونهای شیمیایی در محل اتصال سلولها به همدیگر (سیناپس) با شدت و ضعف مورد نیاز از سلولی به سلول دیگر منتقل میشود. این انتقال سیگنالها در نرونهای مصنوعی توسط مقادیر عددی که به آنها وزن اطلاق میشود انجام میگردد و در نهایت نحوه کارکرد آن نیز مطابق سیستم عصبی جانداران بر اساس یادگیری فرآیند از طریق یک الگو استوار است.

در استفاده از شبکه های عصبی مصنوعی بخش عمده داده ها صرف آموزش و مابقی صرف تست شبکه می گردند . در طی فرایند یادگیری میزان فراگیری شبکه توسط توابع هدف مرتباً سنجیده می شود و در نهایت شبکه ای مورد پذیرش قرار می گیرد که کمترین میزان خطا را دارا می باشد.

در این تحقیق سه نوع مختلف شبکه عصبی مصنوعی که عبارتند از شبکه های پرسپترون چندلایه (Multi – layer perceptron*)، شبکه های برگشتی (* Recurrent*) و شبکه های برگشتی با تاخیر زمانی (*Time lag recurrent*) مورد استفاده قرار گرفته اند.*

لازم به ذکر است که هر سه شبکه عصبی مصنوعی استفاده شده در این تحقیق اساسا جزء شبکه های پیشخور (Feed forward networks) *محسوب میشود. این نوع شبکه ها معمولاً بوسیله الگوریتم انتشار به عقب (*Backpropagation*) آموزش داده میشود.*

٤-سری داده های آماری تحقیق

در این تحقیق ملا ک انتخاب ایستگاههای هیدرومتری جهت آنالیز آمار آن به طور خلاصه به شرح زیر بوده است :

١. آمار ایستگاه قابل اعتماد باشد.

٢. سد مخزنی یا انحرافی در بالا دست ایستگاه انتخابی ایجاد نشده باشد.

٣. ایستگاههای انتخابی از یک پراکنش اقلیمی مناسب در سطح ایران برخوردار باشند.

٤. ایستگاه انتخابی حتی الامکان دارای آمار طولانی مدت باشد (البته در صورتی که شرایط دیگر را دارا باشد.

٥. نواقص آماری ایستگاهها حداقل باشد.

با توجه به موارد فوق حدود ٤٠ ایستگاه هیدرومتری در نقاط مختلف ایران انتخاب و سپس با توجه بیشتر به ملاکهای ذکر شده در نهایت ١٧ عدد از آنها که نسبت به ایستگاههای دیگر با توجه به اهداف این تحقیق برتری داشتند انتخاب گردیده واطلاعات مورد نیاز مربوط به آنها جمع آوری شد. مشخصات این ایستگاهها در جدول ١ نشان داده شده است.

جدول۱ مشخصات ایستگاههای هیدرومتری انتخاب شده

ردیف	نام ایستگاه هیدرومتری	نام رودخانه	موقعیت جغرافیایی ایستگاه		ارتفاع از سطح دریا (متر)	وسعت حوزه بالا دست (کیلومترمربع)	اقلیم حاکم بر حوزه بالادست
			طول	عرض			
۱	پاشاکلا	بابلرود	۵۲-۴۷	۳۶-۱۴	۲۱۲	۱۵۲	نیمه مرطوب معتدل
۲	قران تالار	بابلرود	۵۲-۴۷	۳۶-۱۷	۱۵۰	۴۰۳	نیمه مرطوب معتدل
۳	بابل(کشتارگاه)	بابلرود	۵۲_۳۹	۳۲_۳۶	۰	۱۴۳۰	نیمه مرطوب معتدل
۴	باغستان	ده بالا	۵۴-۰۹	۳۱-۳۹	۲۲۰۰	۷۲	فراخشک سرد
۵	فیض آباد	فیض آباد	۵۴-۰۰	۳۱-۴۳	۲۰۲۰	۱۵۳	فراخشک سرد
۶	طزرجان	طزرجان	۵۴-۰۹	۳۱-۳۹	۲۲۰۰	۲۶	فراخشک سرد
۷	محله پایین	بنادک سادات	۵۴-۱۳	۳۱-۳۴		۲۵	فراخشک سرد
۸	نیر	پشت کوه	۵۴-۰۶	۳۱-۳۰	۲۵۰۰	۳۰	فراخشک سرد
۹	فخرآباد	فخرآباد	۵۴-۱۵	۳۱-۳۸	۱۸۰۰	۲۰۶	فراخشک سرد
۱۰	ایدنک	مارون	۵۰-۲۵	۳۰-۵۷	۵۶۰	۲۷۴۶	خشک و گرم
۱۱	چم نظام	مارون	۴۹-۵۵	۳۰-۴۵	۱۹۰	۵۲۴۰	خشک و گرم
۱۲	تنگ تکاب بهبهان	مارون	۵۰_۲۰	۳۰_۴۱	۲۸۰	۳۷۴۰	خشک و گرم
۱۳	کوشک	کارده	۵۹-۳۶	۳۶-۴۱	۹۸۰	۹۳	نیمه خشک
۱۴	جنگ	کارده	۵۹-۳۵	۳۶-۴۷	۱۶۸۰	۹۵	نیمه خشک
۱۵	کارده بالا	کارده	۵۹-۳۹	۳۶-۴۱	۹۸۰	۴۳۲	نیمه خشک
۱۶	بند بهمن	قره آقاج	۵۲-۳۶	۲۹-۱۳	۱۷۰۰	۲۴۱۰	خشک
۱۷	علی آباد خضر	قره آقاج	۵۳_۳۰	۲۹_۰۰	۱۳۴۰	۳۵۷۰	خشک

۵-روش بررسی

مراحل انجام کار در این تحقیق بطور خلاصه بشرح زیر بوده است:

۱. بررسی منابع و جمع آوری اطلاعات و تحقیقات مرتبط.

۲. بررسی و انتخاب اولیه ایستگاههای هیدرومتری براساس ملاکهاومعیارهای طرح شده دراین خصوص.

۳. بررسی عمیق تر ایستگاههای انتخاب شده از نظر پراکنش اقلیمی ،صحت و طول آمار موجود و انتخاب نهایی ایستگاهها.

٤. اخذ آمار دبی روزانه، ماهانه و لحظه ای ایستگاههای انتخاب شده.

۵. بررسی صحت وهمگنی آمار و پردازش و آماده سازی آن برای مدل شبکه عصبی مصنوعی.

٦. بررسی ومدلسازیهای متعدد شبکه های عصبی مصنوعی جهت رسیدن به شبکه ای مناسب برای هر شبیه سازی.

۷. انجام شبیه سازی براساس داده های آماده شده و مدل شبکه عصبی طراحی شده وتکمیل شبیه ساز یها با ایستگاههای اضافی جهت اطمینان بیشتر.

۸. تجزیه و تحلیل نتایج اخذ شده از مدل شبکه عصبی مصنوعی ومقایسه آن با روشهای سنتی موجود.

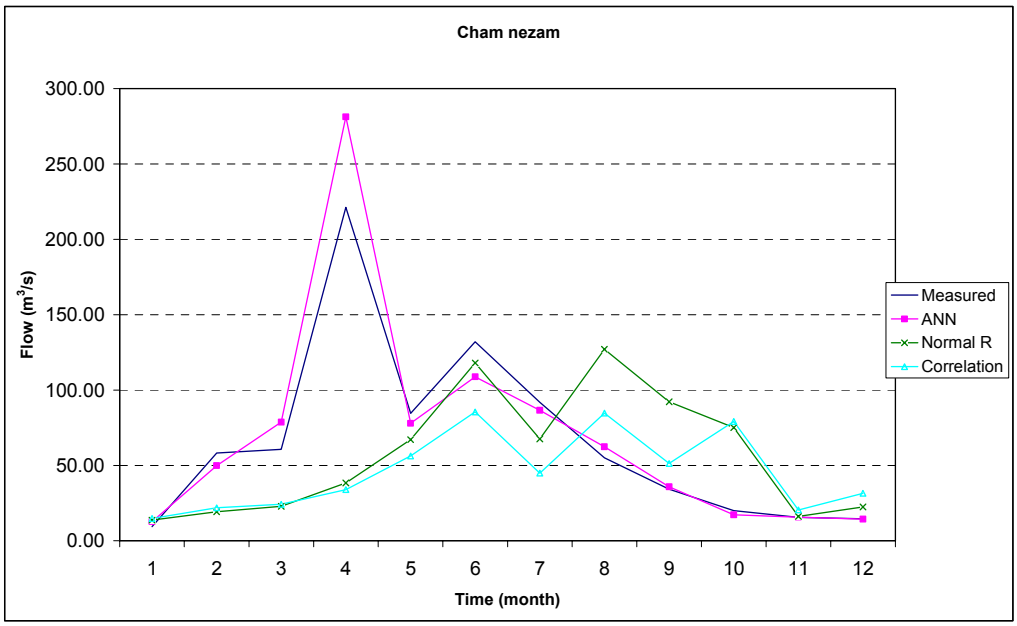
در این تحقیق برای هر یک از ایستگاههای هیدرومتری انتخاب شده دو نوع از داده های اندازه گیری شده آماده سازی و استفاده گردید که عبارت بودند از دبیهای متوسط روزانه و دبیهای متوسط ماهانه. ضمناً ایستگاهها همانطور که در جدول شماره۱ نیز ملاحظه می گردد در قالب گروههایی انتخاب و تنظیم گردید که در کل پنج گروه شامل گروه ایستگاههای بابلرود، مارون، کارده، قره آقاچ و منطقه شیرکوه تنظیم شد.

٦- نتایج بدست آمده

نتایج مربوط به بازسازی مقادیر مفقود شده آمار هیدرولوژیکی توسط شبکه عصبی مصنوعی، روش نسبت نرمال و روش همبستگی به همراه مقادیر واقعی (اندازه گیری شده) مربوط به ایستگاه چم نظام در جدول ۲ درج شده است. علاوه بر آن این مقادیر بصورت گرافیکی در شکل ۱ مقایسه شده است.

جدول ۲ .مقادیر برآورد شده توسط روشهای مختلف در مقایسه با مقادیر واقعی اندازه گیری شده در ایستگاه هیدرومتری چم نظام (رودخانه مارون).

روش همبستگی	روش نسبت نرمال	شبکه عصبی مصنوعی	مقادیر اندازه گیری شده
۱٤,۸۰	۱۳,۷۱	۱۲,۷٤	۹,٤۳
۲۱,۹۵	۱۹,۲۰	٤۹,۸٤	۵۸,۳۳
۲٤,۲۱	۲۲,۸۵	۷۸,٦۸	٦۰,٦۸
۳۳,۹٤	۳۸,۳٤	۲۸۱,۳۳	۲۲۱,۳۰
۵٦,۳۲	٦۷,۱۰	۷۷,۹۳	۸٤,۵۸
۸۵,۵۰	۱۱۸,۰۱	۱۰۸,۸۱	۱۳۲,۰۲
٤٤,۹۱	٦۷,۵۵	۸٦,٦۰	۹۱,۹٤
۸٤,٦۷	۱۲۷,۰۷	٦۲,٤۲	۵۵,۰۹
۵۱,۲۷	۹۲,۱۷	۳۵,۷۸	۳٤,۳۳
۷۹,۲۳	۷۵,۱۲	۱۷,۲۳	۲۰,۰۲
۲۰,٤٤	۱٦,۱۸	۱۵,۷۱	۱۵,۵۸
۳۱,۵۱	۲۲,٤۳	۱٤,٤۲	۱٤,۵٦



شکل ۱.هیدروگراف مقادیر برآورد شده توسط روشهای مختلف در مقایسه با مقادیر واقعی اندازه گیری شده در ایستگاه چم نظام (رودخانه مارون).

جدول ۳. مقدار مربوط به ریشه میانگین مربع خطا و ضریب کارایی برای مقادیر برآورده شده با روشهای مختلف ومقادیر واقعی برای ایستگاههای مختلف استفاده شده در این تحقیق

ایستگاه هیدرومتری	R^2			$RMSE$		
	شبکه عصبی	نسبت نرمال	همبستگی	شبکه عصبی	نسبت نرمال	همبستگی
چم نظام	۰,۹۵۰	۰,٤٧٢	۰,٤٨٥	۱۹,۷٦۱	٦٤,۰۱۷	٦۳,۲۰٥
فخر آباد	۰,۷۸۳	۰,۳۲٥	۰,۸۷۳	۰,٤۷۰	۰,۸۳۰	۰,۳٦۰
بابل	۰,۸٤۱	۰,۸۷۹	۰,۸۱۳	٦,٦۷٤	٥,۸۰٦	۷,۲۲۲
علی آباد	۰,۸۰۰	۰,٤٤۰	۰,٦٥٤	۲,۳۹٦	٤,۰۱۲	۳,۱٥۹
کارده	۰,۸۲۱	۰,٥۳۷	۰,۲۱۰	۰,۱٤۰	۰,۲۲٥	۰,۳٦٥

۷-نتیجه گیری نهایی وپیشنهادات

این تحقیق در حقیقت سعی درارزیابی کاربرد جدیدی از شبکه های عصبی مصنوعی را که همانا توانایی آن در بازسازی نواقص آماری سری های جریان در نقاط مختلف ایران و با شرایط اقلیمی متفاوت است داشته .

جهت امکان مقایسه نتایج حاصل از شبکه های عصبی مصنوعی با روشهای موجود، روش نسبت نرمال و روش همبستگی بین ایستگاهها نیز به کار گرفته شده و مقادیر برآورده شده شبکه های عصبی مصنوعی با این روشها نیز برآورد گردید.

بر اساس نتایج به دست آمده از این تحقیق میتوان گفت که تکنیک شبکه های عصبی در حقیقت کارایی مناسب جهت پیش بینی داده های مفقود شده هیدرولوژی را دارد. باتوجه به مقادیر مربوط به ریشه میانگین مربع خطا ($RMSE$) و ضریب کارایی(R^2) برای نتایج ارائه شده توسط هریک از روشها و مقادیر واقعی (اندازه گیری شده)، در بیشتر موارد در این تحقیق درستی نتایج حاصل از شبکه های عصبی مصنوعی بالاتر از نتایج بدست آمده از دو روش دیگر بوده است . شبکه عصبی مصنوعی از نوع پرسپترون چند لایه سازگار ترین ساختار برای این هدف بوده است هرچند که در موارد معدودی شبکه عصبی برگشتی نیز نتایجی همانند شبکه پرسپترون چند لایه ویا حتی اندکی بهتر از آنرا ارائه نموده است .

علاوه بر تناسب و سازگاری نوع شبکه عصبی مصنوعی برای این مسئله، صحت و اعتماد به داده های مورد استفاده جهت آموزش مدل جهت رسیدن به نتایج قابل قبول بسیار مهم تشخیص داده شد. با انتخاب ایستگاههای شاهد مناسب و مطمئن و نیز استفاده از نوع مناسب و سازگار شبکه عصبی مصنوعی و نیز کالیبره کردن مناسب آن میتوان گفت که این تکنیک ابزار بسیار کارآ و مناسبی برای حل مشکل برآورد مقادیر مفقود شده و ناقص در سریهای هیدرولوژیکی است. بدیهی است جهت اظهار جامع و کاملتر در خصوص روش مورد اشتفاده در این تحقیق و مقایسه با روشهای دیگر نیاز به تحقیقات بیشتر با استفاده از داده های مناطق دیگر کشور ضروری است.

۸-فهرست منابع

۱- مهدوی، محمد، هیدرولوژی کاربردی، انتشارات دانشگاه تهران ۱۳۷۱.

۲- Dastorani, Mohammad T. ۲۰۰۲. Implementation and evaluation of artificial neural networks for river flood prediction, Ph PhD Thesis. University of Nottingham, Nottingham. UK.

۳- Dastorani, M.T and N. G. Wright, Artificial Neural Network Based Real-time Flow Prediction, ۵th International Conference on Hydrodynamics, July ۱-۵, ۲۰۰۲, Cardiff, UK.

٤- Dastorani, Mohammad T. & Nigel G. Wright, Artificial neural network based real-time river flow prediction, Proceedings of the Fifth international conference of Hydrodynamics, July ۱-۵ ۲۰۰۲, Cardiff, UK.

۵-Wright N. G. and M. T. Dastorani, Effects of river basin classification on Artificial Neural Networks based ungauged catchment flood prediction, in the Proceedings of the International Symposium on Environmental Hydraulics, December ۵-۸, ۲۰۰۱, Phoenix, USA

٦-Wright N.G., M. T. Dastorani, P. Goodwin and C. W. Slaughter, Using artificial neural networks for optimisation of hydraulic river flow modelling results, Proceedings of the International Conference of River Flow ۲۰۰۲, September ٤-٦, ۲۰۰۲, Louvain-al-Neuve, Belgium.