Fuzzy mapping of momentum fluxes in complex shear flows with limited data
B. G. Hankin, H. Kettle and K. J. Beven

ABSTRACT
This study is divided into three parts centred around modelling the complex turbulent fluxes across strong shear layers, such as exist between the channel and floodplain flow in an over bank flood flow. The three stages utilize Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to make fuzzy mappings between the fluxes and different data types. The de-fuzzification stage commonly used in Fuzzy Inference Systems is adapted to avoid the generation of crisp outputs, a process which tends to hide the underlying uncertainty implicit in the fuzzy relationship.

Each stage of the study utilizes conditioning data that makes the fuzzy mappings more tenuously linked with what would normally be considered physically based relationships. The need to make such mappings in distributed models of complex systems, such as flood models, stems from the sparsity of available distributed information (e.g. roughness) with which to condition the models. If patterns in distributed observables which clearly affect, or are affected by, the river hydraulics can be linked to the local fluxes, then the conditioning of the model would improve. Mappings such as these often suffer from scaling effects, an issue addressed here through training the fuzzy rules on the basis of both laboratory and field collected data.

Key words | de-fuzzification, fuzzy logic, neural network, scaling

INTRODUCTION
Accurate representation of the partitioning of momentum flux between a river channel and floodplain is crucial to predictions of inundation extent during a large event (Bates et al. 1996; Knight & Shiono 1996; Aronica et al. 1998). The information necessary to drive a complex distributed model of such an event is subject to a large degree of non-random uncertainty, or fuzziness, reflecting a lack of knowledge about the key variables that control the predicted flow. For large scale events, the controls are widely held to be distributed bed slope and effective roughness (Romanowicz et al. 1996; Aronica et al. 1998). The extent to which each of these variables, coupled with antecedent flow conditions, influence the diffusive and convective transport properties of the flow is unclear, and there are unlikely ever to be complete data sets available with which to validate predictions of distributed flood models at the large scale. The turbulent fluxes of mass and momentum across transverse shear layers within an engineered flume flow at the Flood Channel Facility (FCF) have been studied in detail (Knight & Shiono 1989), from which we have a better understanding of the effect that the shear has on the overall dynamics, and the resulting dispersion characteristics of the flow (Hankin & Beven 1998a, b). Yet the dynamics of such phenomena in real systems, with far more complex boundary conditions, for which we have sparse distributed information, is far from understood (Bates et al. 1996). Characterizing the turbulent momentum fluxes across a strong shear layer therefore forms the main area of interest in this study.

It has already been shown that the traditional numerical representation of the partial differential equation (PDE) for pure advection (Minns 1998) or the wave equation (Dibike et al. 1999) or for two-dimensional flow (Dibike & Abbott 1999) can equally well be represented in
terms of the weights in an artificial neural network (Babovic & Abbott 1997). In this approach, artificial neural networks (ANNs), trained using numerical representations of the PDE for a particular problem, are shown to reinstate the governing PDEs, and it is argued that they must therefore contain the same knowledge, or have the same content, as the traditionally favoured continuum equations. If this is the case, then data taken from nature should also provide ANNs that can be used to produce new PDEs. For a complex river or flood flow, the data is often sparse, and distributed flow information is extremely difficult to collect, so in this study Adaptive Neural Networks based on Fuzzy Inference Systems, or ANFIS, (Jang & Gulley 1995) are used. By training ANNs based around fuzzy relationships, we sacrifice the ability to make precise data based statements, but recognize that there is a substantial degree of non-random uncertainty in what data we might have accrued. The kinds of models under development for flood flows by the authors are based around finite volume approaches (Versteeg & Malalasekera 1995), so it is natural to develop rules for the discretized fluxes between control volumes in order to have sufficient information to drive the flow, making the derivation of the PDEs unnecessary. The process of building up fuzzy rules for flux partitioning as a result of complex shear layers are detailed below in each of three sections.

The first application involved using detailed Acoustic Doppler Velocimeter (ADV) measurements in and around an artificial dead zone, so that fluxes in the region of shear between the main flow and the dead zone could be studied. The artificial dead zone was placed in two small flumes with different flows, in order to try to generate scale independent fuzzy rules using ANFIS between the local velocity gradients and the local Reynolds stresses, which can be interpreted as turbulent fluxes. Reasonable agreement between the predictions for these two flows was found, so the ANFIS mapping was applied to estimation of turbulent fluxes in a large natural fluvial dead zone.

The second part of this study addresses a similarly complex shear zone above the bank top in an overbank flow (Knight & Shiono 1989). This time we assume that in an even more data-sparse environment, the only conditioning on the turbulent partitioning that we have is the local bathymetry, or geometry. It is speculated that the boundary conditions are largely responsible for forcing the turbulent motion of this kind of flow, preventing the turbulence from becoming completely chaotic, so it is natural to consider the local topography as having strong influences on the flow itself. The strong fits to the data using ANFIS were then used in the recall mode (when an ANN is applied to data not used in training) successfully to predict some measurements of turbulent fluxes in a real flood flow on the River Severn, UK, to within prediction uncertainty bounds.

The third part of the study is less specific, but still related to shear zones, and the impetus for it stems from the efficiency gains that could be made in large scale river studies, if remotely sensed thermal images (see Figure 1 (after Reynolds et al. 1991)) could be used directly to glean
spatial information about the hydraulics, which have in part generated the observed temperature distribution. We see that in regions of quiescent flow, such as in the dead zone, or along the river banks, that the water surface temperature is warmer.

METHODS

For a more detailed overview of fuzzy sets, readers are referred to texts with environmental applications such as Bardossy & Duckstein (1995) and Ross (1995). The following describes fuzzy sets and rules which are detailed schematically in Figure 2.

As an example fuzzy set, let \( dU/dy \) be the set (universe) of observed velocity gradients anywhere in a flow field. Let \((dU/dy)_{small}\) be a subset of ‘small’ velocity gradients. \((dU/dy)_{small}\) is a set of ordered pairs:

\[
(dU/dy)_{small} = \{ (y, \mu_{(dU/dy)_{small}}(y)) ; \ y \in dU/dy, \ \mu_{(dU/dy)_{small}}(y) \in [0,1] \}
\]

where \( \mu_{(dU/dy)_{small}} \) is the grade of membership of \( y \), the local velocity gradient, in \((dU/dy)_{small}\), and is termed the membership function of \((dU/dy)_{small}\). A membership of 0 or 1 implies \( y \) absolutely does not or does belong to \((dU/dy)_{small}\) respectively. Those membership grades between imply that \( y \) has partial membership of \((dU/dy)_{small}\). Three fuzzy subsets of \( dU/dy \) (see Figure 2) are used to describe the complete set velocity gradients, given by ‘small’, ‘medium’ and ‘large’, each having membership functions with a certain shape such as a simple triangular or trapezoidal function. Each shape is described by a set of simple shape parameters, and the extent or range over which a fuzzy set is non-zero is called the support.

The general form of a fuzzy rule (Figure 2) consists of a set of premises, \( A_{i,f} \), in the form of fuzzy sets (numbers) with membership functions \( \mu_{A_{i,f}} \) and a consequent, \( L_i \) (in this study this always refer to a turbulent flux), also in the form of a fuzzy set:

\[
\text{IF } A_{i,1} \odot A_{i,2} \odot \ldots \odot A_{i,f} \text{ THEN } L_i
\]

where \( i \) refers to the rule number and \( f \) to the piece of information the relation pertains to, for example in Equation (2), rule number 1, \( A_{1,2} \) might refer to the premise: ‘if the transverse velocity gradient is large then the turbulent flux is large’. The operator \( \odot \) is the fuzzy logic equivalent of the Boolean logic AND or OR statement which is used to combine premises based on different pieces of information. For a particular rule, the fuzzy operator for ‘OR’ is typically the MAX (maximum) of the scores that a set of inputs have in each premise fuzzy set that is involved in the rule (Figure 2). Likewise the AND operator is typically the MIN (minimum) of these scores (Ross 1995). The THEN part of the rule is put in effect by an implication operator which is often a simple truncation operation performed on the consequent fuzzy set (Figure 2). Once all of the consequent fuzzy sets have been acted on by all of the fuzzy rules, the truncated consequents must be combined in a process termed aggregation, usually via another MAX operation. The result of implication and then aggregation is a consequent fuzzy set, which typically gets converted into a single crisp output through taking the weighted average of the outputs (Jang & Gulley 1995).

In this study the Adaptive Neuro-Fuzzy Inference System (ANFIS) model of Jang (1993) is used, in which Sugeno-style fuzzy rules are used (Jang & Gulley 1995). Sugeno-style fuzzy rules link the premises of each rule to a consequent, which is in the form of the equation of a line rather than a fuzzy set. Such an approach leads to more computationally efficient rules, but has less intuitive outputs (Jang & Gulley 1995), represented in Figure 2 by lines which can move around the consequent universe.

The shape parameters of the input fuzzy sets and the coefficients of the lines of the consequents are then optimized through the use of an ANN, based on the fit that the resulting rules give to a set of training data.

The number of fuzzy subsets for each input and output universe can be chosen quantitatively through gradually increasing the number of fuzzy subsets, and observing the change to a measure of goodness of fit that the ANFIS achieves in reproducing the training data. If relatively little data is available then the goodness of fit will not improve beyond a few fuzzy subsets, resulting in mappings with a relatively low information content. This will normally be the case, because if there is an abundance of data with a high information content, then there is little need to
Figure 2 | An example Sugeno-style fuzzy inference system.
use fuzzy techniques. Clearly, there is an upper bound to the number of subsets that can be used, or the situation could result in which there are more variable parameters in the ANN that is being calibrated, than there are items of data. There is also the problem of over-training the rules, such that they form an exceptionally good fit to the training data, but are not very transferable to other data sets. By reducing the continuum of modelled turbulent fluxes into three sets, for example, we can then incorporate our limited number of fuzzy observations relating fluxes to local velocity gradients.

In this study an attempt is also made to retain the richness of ‘non-de-fuzzified’ consequent (see the lower option at the very base of Figure 2) which reflects the uncertainty in all of the rules. There are a number of techniques for propagating fuzziness where the mapping is analytical (Dou et al. 1997) based around the extension principle (see Zadeh 1975), but the simplest approach for FIS mappings appears to be to change the de-fuzzification stage which is normally used to produce a crisp number from the consequent fuzzy sets. This is achieved here through taking a ‘fuzzy’ measure, expressed using the deviation of each weighted fuzzy consequent away from the weighted average. The weighted average plus or minus this mean deviation is then used to provide prediction limits. The resulting prediction limits should be interpreted as a fuzzy envelope of possible mappings. A possibilistic, as opposed to probabilistic, framework has been used since the limits are intended to indicate a gap in our knowledge, which includes knowledge about the distribution of errors. The approach is not very robust for cases where there are few rules, and hence few outputs from which to take the deviation measure. Nonetheless it represents an attempt to move away from treating Fuzzy Inference Systems in complex systems as another tool for deterministic modelling.

RESULTS

Stage 1: Mapping fluxes in the shear generated by flow past dead zones

Fluxes across a region of shear produced specifically by a dead zone have been modelled using computational fluid dynamics (CFD) by El Latif & Campbell (1995) and Kimura (1997). Dead zone features (slow moving or recirculating regions of water) are important to the longitudinal dispersion characteristics of a complex river flow, giving rise to the long tail in concentration break-through curves, as highlighted in a number of studies (Wallis et al. 1989; Heslop & Allen 1993; Young & Wallis 1993; Rutherford 1994). The processes governing fluxes across strong transverse shear layers affect both solute and suspended sediment dynamics (Tipping et al. 1989; Carling et al. 1994), and the resilience of some ecological systems (RSPB 1994).

Figure 3 represents the geometry of the set-up of two flume flows in which artificial dead zone features have been added, characterized by:

1. Large flume: 0.895 m wide, 0.24 m deep, mean velocity 0.1 m/s.
2. Small flume: 0.305 m wide, 0.28 m deep, mean velocity 0.5 m/s.

The left hand side sub-plots in Figure 4 show plan views of transverse velocity gradients plotted against the y (transverse) locations across the shear zone (see Figure 3).
These are used as premises to construct fuzzy inference system of the form given in Equation (3):

If $dU/dy$ is small then $-Ruv$ is small
If $dU/dy$ is medium then $-Ruv$ is medium
If $dU/dy$ is large then $-Ruv$ is large

where $-Ruv$ is the transverse cross correlation between the turbulent transverse and downstream velocity components, also termed the transverse Reynolds stress, although we have omitted to multiply it by the density in order to make it dimensionally homogenous to a stress, for simplicity.

Sets of data from both flumes were used to construct an ANFIS based optimal fuzzy mapping, in order to try and generate a set of rules that would be more independent of the flow and geometry. At this stage, we are not interested in the efficiency of using the mappings in the recall sense, as we are trying to make a scale-independent mapping. From the right hand side of Figure 4, this is clearly possible, with good fits to two flows of different depths, discharges and dimensions. The prediction limits

![Graphs showing observed dU/dy and -Ruv (o) and FIS predictions (-) / prediction limits (-.) of -Ruv for large flume and small flume](https://iwaponline.com/jh/article-pdf/3/2/91/392228/91.pdf)
show how the fuzziness in the inference approach is represented at the prediction stage, as discussed above.

**Application to measurements in a real dead zone**

When the same ANFIS mapping was used in the recall sense to velocity and turbulent velocity data from electromagnetic current meter (ECM) measurements around a dead zone on the River Severn (Figures 1 and 5, showing a CFD generated flow around the dead zone), a poor fit was obtained. This was due to the extrapolation problem: there were scales of velocity gradients that simply had not been encountered in the training of the fuzzy rules. For this reason, the River Severn data was then used in addition to the flume data as more training data, and the new fuzzy mapping gave the fit shown in Figure 6. This is still regarded as a poor fit, although with more training data, the universality of the rules that was found between the two different scale flume flows might yet be achieved for large scale flows. As more data becomes available it may become apparent that in order to allow for more universal mappings, further rules that reflect more general antecedent conditions must be used, such as ‘if the Reynolds Number (indicating scale of flow) is large then use rule set number two’ and ‘if the Reynolds Number is small then use rule set one’. The recall ANFIS for the next application was more successful.
Stage 2: Mapping fluxes across the shear between channel and floodplain flow

The experimental data used for training here was collected in the 50 m long Flood Channel Facility at Hydraulics Research, Wallingford, in a two stage, straight channel geometry, with a main channel depth of 176.5 mm and a side wall slope of 2, as shown in Figure 7. The turbulence and velocity data were collected using a laser Doppler anemometer with a spatial resolution of 2 mm in the transverse direction and 0.5 mm in the vertical direction at a sampling frequency between 20 and 100 Hz (Knight & Shiono 1989). The measurements were limited in the overbank region since the transverse and vertical mean velocities were available only at a single depth.

A new set of fuzzy rules were constructed using the premises of local relative depth ($R_d =$ measurement depth/total local depth) and local distance to bank-top (local $y$ coordinate separation from bank top, non-dimensionalized using main channel depth), since these would be topographical features that could be readily obtained from real river systems. Half of the data was used to train the fuzzy rules, and the rest was used to check the efficiency of the fit.

The closeness of fit was measured using the coefficient of determination, $C_D$, given by:

$$C_D = 1 - \left( \frac{\sum_{i=1}^{N_F} (F_{Pi} - F_{Oi})^2/N_F}{s_{data}^2} \right)$$

where the summation is over the number of data points, $N_F$. $F_{Pi}$ and $F_{Oi}$ are the predicted and measured fluxes respectively at site $i$, and $s_{data}$ is the standard deviation for the data. Figure 8 indicates that very close mappings of fluxes through the depths are possible with such detailed measurements, with both vertical and transverse fluxes (training and checking data combined) being predicted with coefficients of determination around 0.8. However, the number of parameters involved in the training of the FIS was large (12), so the goodness of fit is perhaps not surprising.

Application to measurements in a real over bank flow

The above ANFIS was next applied in a recall sense to unpublished ADV measurements from the River Severn in flood, and the closeness of fit accounting for prediction uncertainties, was found to be good, as shown in Figure 9. Although there is not much data to compare such rules with, the mapping would appear to imply that, despite the topographically rich field environment, there are similar gross flow processes occurring as compared with the flume situation. Fuzzy relationships may be the most realistic way of representing these features, despite advances in turbulence modelling, because we simply do not have the detailed boundary information necessary to make more accurate process representations for the kind of large scale systems that this study is aiming towards.

Stage 3: Linking remotely sensed information to hydraulics

This application builds on an earlier study (Hankin et al. 2000 in review) in which the depth-averaged hydraulics of a natural fluvial dead zone were modelled using a finite element based St Venant equation solver, TELEMAC-2D (Hervouet & Van Haren 1995), with semi-distributed bed roughness and eddy viscosities in three distinct zones of the dead zone, shear zone and main channel. The constant eddy viscosity assumption led to a relatively simple hydraulic model, with few claims to represent the precise turbulent behaviour of the flow, but it was found to retain sufficient flexibility to capture the functionality of the system in terms of the observations that were available. Indeed, we believe that there are a whole host of models and model structures which would achieve this, giving rise to the phenomenon of equifinality (Beven & Binley 1992; Beven 1995; Aronica et al. 1998) of models for complex...
systems. Using net energy budget data from measurements in and above the dead zone, and assuming that buoyancy forces could be neglected, this earlier study indicated that the gross hydraulics of the system appeared to explain much of the temperature distribution that was observed using infra-red imagery, collected using the NERC Daedalus AADS thermal line scanner (Carling et al. 1994). With this in mind, it was conjectured that a remotely sensed thermal image of the dead zone could be used directly to glean spatial information about the hydraulics which have in part generated the observed temperature distribution.

In the CFD scheme, the conservative streamline upwind Petrov–Galerkin (SUPG) numerical scheme was used to model the advection of depth in the continuity equation, forcing conservation of mass everywhere except at the downstream boundary, where the free surface was imposed. An implicit N scheme, which is unconditionally stable, was used for the advective velocities. A one second time step was used, in order to maintain a Courant number of approximately unity, and the model was run for 10,000 and 20,000 time steps, the latter producing the vector plot given in Figure 5. The two simulations yielded a whole field residual (between the above successive

Figure 8 | Prediction of the turbulent fluxes based on data from the Flood Channel Facility using ANFIS generated rules based on local geometry.
computations) of 0.0009 m/s and 0.002 m/s for down-
stream ($U$) and cross-stream ($V$) velocities respectively.

The closeness of fit to the depth averaged mean vel-
ocity measurements (of which there were 30 distributed
around the 3 zones) gave a coefficient of determination of
0.53. This was considered to be a sufficiently good numerical
simulation of the system on which to build the fuzzy
rules. The next objective was to make a set of fuzzy
inferences about the relationship between the gross
hydraulics and the surface temperature distribution using
information that can be gleaned from over-flight data
alone. This restriction on the set of premises, or anteced-
ent conditions, led to the choice of premises of remotely
sensed temperature, temperature gradient and channel
width, which are all believed to affect or be affected by the
flow. As an example, it could be assumed that at the scale
of interest (several metres) there is a relationship between
the strong temperature gradient across the shear zone and
the advective momentum transport across it, which can be
expressed as a fuzzy linguistic rule:

If [transverse temperature gradient] is large then [trans-
verse advective flux] is small \( (5) \)

where the italics refer to fuzzy sets of quantities that are
not known precisely. Each of 3 premises (temperature,
temperature gradient and local channel width), were split
into 2 ranges representing a small or a large quantity.

There appeared to be no advantage to using three fuzzy
subsets, in that the overall measure of closeness of fit to
the training data did not improve significantly, implying
that there is a limited information content in the data set.

These were linked to the 2 ranges of the consequent, that
is a small or large {advective flux}. Two sets of rules were
built up in this way between the three premises above and
the TELEMAC-2D predicted transverse advective velocity
($V$) and downstream advective velocity ($U$) for every cell
of an experimental finite volume grid. The shape par-
parameters which define the fuzzy sets were then allowed to
vary using the ANFIS tool once more, and optimized to fit
half of the data. The other half of the data was used as
checking data, since the degree of goodness of fit for
any set of rules needs to be offset against the universality
of the resulting set of fuzzy rules. Both velocity com-
ponents were fitted to the training and checking data
combined with an overall coefficient of determination of
approximately 0.6.

Figures 10(a) and (b) indicate the spatial distribution
of the data for the temperature and temperature gradient
premises (the channel width at each location is also
inferred from these plots). Figures 10(c) and (d) indicate
the TELEMAC-2D downstream and cross-stream vel-
ocities assumed to be representative of the flow for the low
stage flow on the day of the over-flight. Half of this data
was used in training the fuzzy rules. The optimized rules were used to produce Figures 10(e) and (f), for which the scales are identical, so the qualitative agreement between actual and fuzzy predicted velocity distributions can be compared. The velocity components were then combined to produce the vector plots of TELEMAC-2D velocities, and fuzzy vector velocities in Figures 10(g) and (h). The fuzzy rule system captures much of the ‘observed’ velocity distribution, although the most difficult velocities to resolve occur in the shear and areas of recirculating flow.

Although only half of the data was used to train the data for this application, the portion used in training was taken from cells throughout the reach, since each part of the reach contained information too disparate to make universal rules. Some new remotely sensed imagery is under investigation, which will be used to check the universality of the rules. Such information would be extremely useful for identifying approximately where there might be zones of pollutant retention, or good mixing, at an accuracy that is perhaps compatible with, for instance, one-dimensional advection dispersion modelling.

For all three applications, we have attempted to make the fuzzy mappings of momentum transport across regions
of strong shear more universal by selecting premises with enough flexibility to reflect the diversity of antecedent flow conditions. It would appear that transferability of the mappings between different scale flows is better when the premises in the rules are selected to be quantities that have a strong influence on the large scale flow features. For instance, the relative depth of flow used in the second application is a good example of how a premise can be selected in order to anticipate that turbulent fluxes across a shear zone will vary between different Reynolds number flows. For flows that are relatively shallow, the rules associated with large depths and stronger momentum exchanges will only weakly influence the ANFIS. As more data becomes available with which to train the ANFIS, so the mappings for different scale flows will improve.

The fuzzy logic framework allows us to make broad statements about different scale flows, based on sparse data. Such approximate mappings will be useful in the design of the weir type relations for transfer of channel flow-to-storage zones in the 1D and 2D simplified flood flow models that are widely used.

CONCLUSIONS

1. A technique which avoids the complete de-fuzzification part of fuzzy inference modelling has been suggested as a way of propagating the uncertainties in the inputs to the predictions. It was argued that fuzzy rule based approaches should not be used to produce the ‘only’ or optimum model structure for complex systems, where there is a large degree of fuzziness.

2. A scale invariant fuzzy mapping was made between mean velocity gradients and turbulent fluxes across strong shear in two flume flows, which could be used as an alternative turbulent closure to the flow equations in complex flows, where process representation is uncertain. It appears that this mapping could be extended to larger flows with the collection of more data.

3. A fuzzy mapping of the turbulent fluxes between the main channel and floodplain was made from data from the Flood Channel Facility and was successfully used in recall mode to estimate the same fluxes in a real flood flow. The importance of obtaining a greater understanding of these interactions for flood flow modelling cannot be stressed enough, and has been widely called for by practitioners. The fuzzy relationships could be used as modules within more complex CFD modelling, and replace for instance a k−ε type closure to the Reynolds averaged Navier–Stokes equation.

4. The fuzzy dependence of the gross hydraulics of a river flow on the water surface temperature distribution taken from a remotely sensed platform has been investigated, producing a relatively successful mapping which could aid in the rapid reconnaissance and classification of large river systems in terms of their mixing characteristics.

ACKNOWLEDGEMENTS

This work has been supported by NERC grant GR3/10687. The authors wish to express their gratitude to Walt Holland for collection of field data, and to all his assistants. The thermal image of Figure 1 was obtained during the 1991 NERC Airborne Remote Sensing Campaign.

REFERENCES


