# Quantitative comparison of spatial fields


## Stephen Wealands[a b], Rodger Grayson[a b], Jeffrey Walker[a] and Günter Blöschl[c]


## OVERVIEW

Comparing spatial model simulations against spatial observations is useful throughout the modelling process (e.g. calibration, validation). Most similarity measures currently in use summarise only the general features of a spatial field. These ignore the spatial arrangement that is of particular interest in many hydrological applications (e.g. wet areas around a stream). The only comparison method that currently utilises the arrangement of values is visual comparison, but it is neither quantitative nor repeatable.

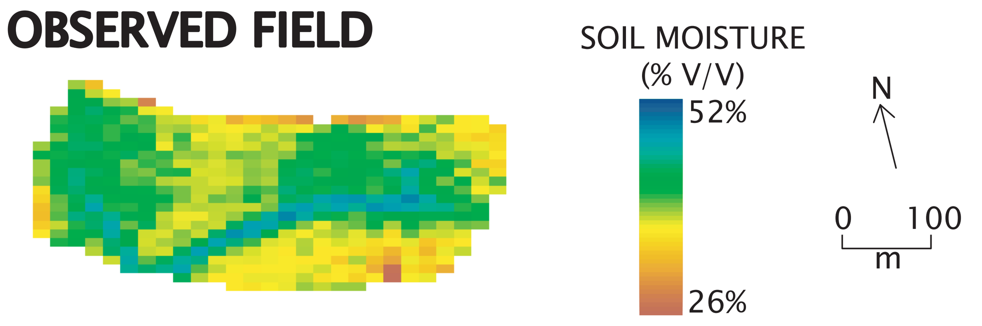
Our aim is to assess and develop similarity measures that describe various aspects of spatial similarity (by using the arrangement of values). We are focussing on local similarity measures, which compare each spatial element, producing a similarity field that can be inspected and then summarised into a numerical similarity measure.

The four methods presented emulate particular aspects of visual comparison.

1. **Fuzzy similarity** tolerates shifts/differences between the fields
2. **Segmentation** recognises regions within a field
3. **Weighted similarity** focuses on important parts of the fields
4. **Multiscale similarity** compares the fields at multiple coarser scales

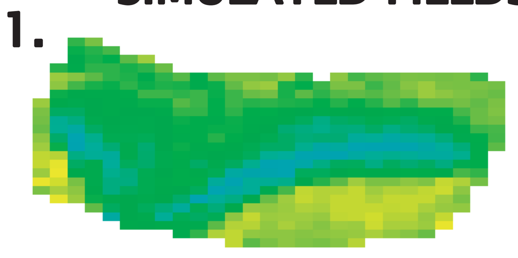
## COMMON MEASURES

### OBSERVED FIELD



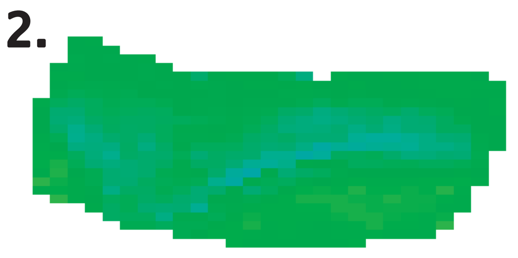

### SIMULATED FIELDS

1.



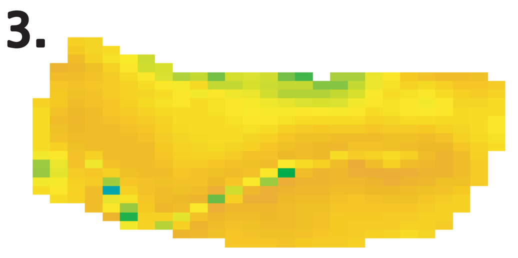
BIAS = 2.0, RMSE = 3.6, MAE = 2.9
RSQ = 0.48, COE = 0.24

2.



BIAS = 3.6%, RMSE = 5.0, MAE = 4.2
RSQ = 0.33, COE = −0.51

3.



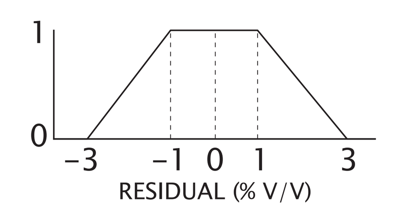
BIAS = −5.7, RMSE = 7.4, MAE = 6.3
RSQ = 0.01, COE = −2.33

**Mean error** (BIAS) measures the global over/under prediction of a simulation. **Root mean squared error** (RMSE) measures the typical error. **Mean absolute error** (MAE) also measures the typical error, but is less influenced by extreme residuals.

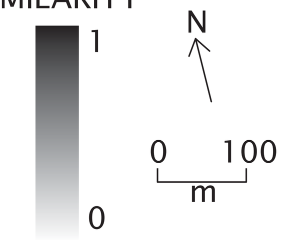
The **coefficient of determination** (RSQ) measures the linear fit between the simulation and observation. The **coefficient of efficiency** (COE) measures how good a predictor the simulation is compared to the observed mean (i.e. with no variation). All these measures ignore the spatial arrangement of the values.

## FUZZY SIMILARITY

MEMBERSHIPS FOR RESIDUALS & LOCATION

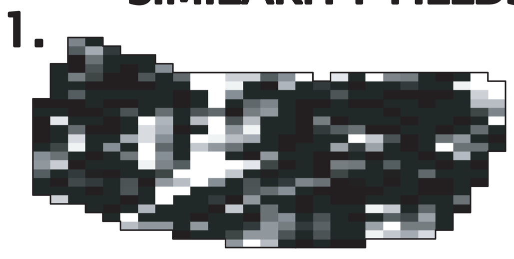



| | | |
|---|---|---|
| 0 | 0.9 | 0 |
| 0.9 | 1 | 0.9 |
| 0 | 0.9 | 0 |



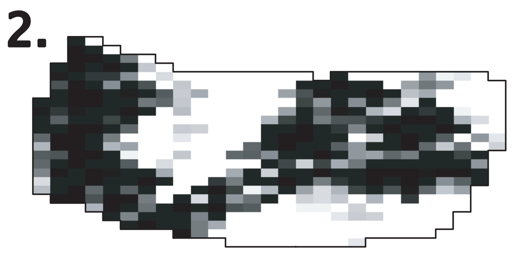

### SIMILARITY FIELDS

1.



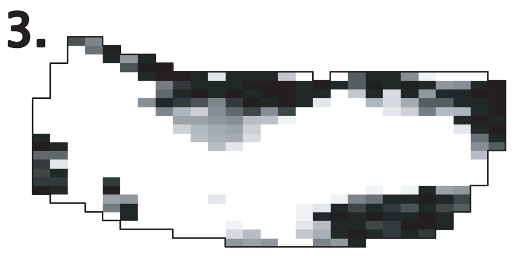
FS = 0.74, FE = 0.11 *

2.



FS = 0.49, FE = −0.71 *

3.



FS = 0.28, FE = −1.42 *

* OBS MEAN: FS = 0.70, FE = 0

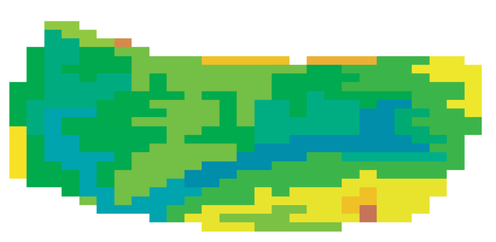
During visual comparison, we accept differences in value and location. This process can be included in a comparison using fuzzy set theory.

**Fuzzy similarity** (FS) measures the average similarity between two fields based on the tolerances specified. If elements differ by less than the tolerances, they are judged as being similar (in the range 0 to 1).

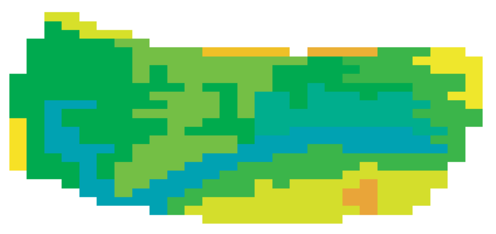
**Fuzzy efficiency** (FE) uses the similarity found with the observed mean to adjust the fuzzy similarity value. A positive value indicates that the simulation is a better predictor than the observed mean (i.e. no variability).

## SEGMENTATION

### SEGMENTED FIELD



SCALE = 200



SCALE = 400

SCALE = 1000

Recognition of regions within fields is innate to humans, but it is not utilised in any similarity measures. **Segmentation** is the process of delineating regions from a field of values. We use a region merging process to define regions that have connected elements and similar values (i.e. minimise increase of variance during each merge). A scale parameter limits the size of the final regions.

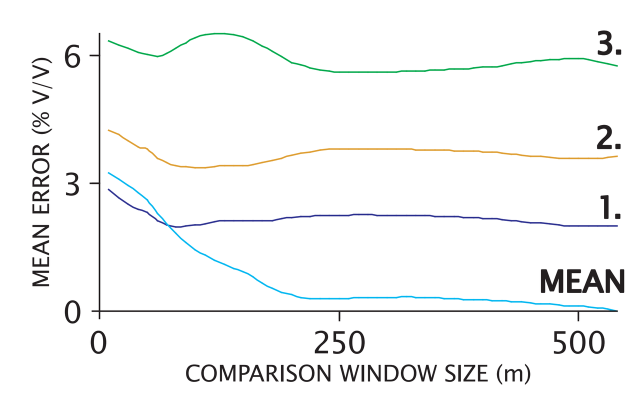
Regions are used for specialised comparisons of spatial fields. They provide logical areas that can be used for weighting comparison measures. They are also multiscale representations of the field.

THE OBSERVED FIELD IS SEGMENTED USING REGION MERGING, WITH THREE DIFFERENT SCALE PARAMETERS.

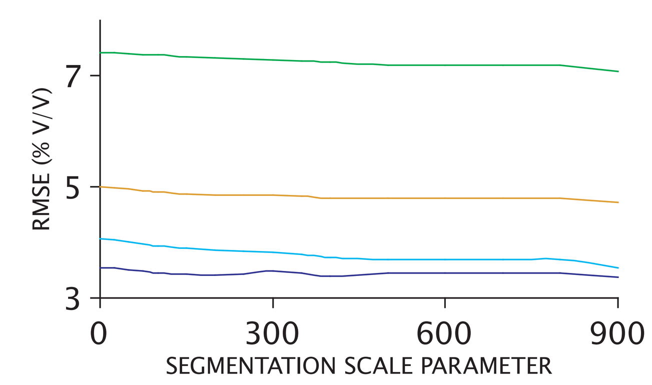
THE COLOURS REPRESENT THE MEAN VALUES FOR EACH REGION.

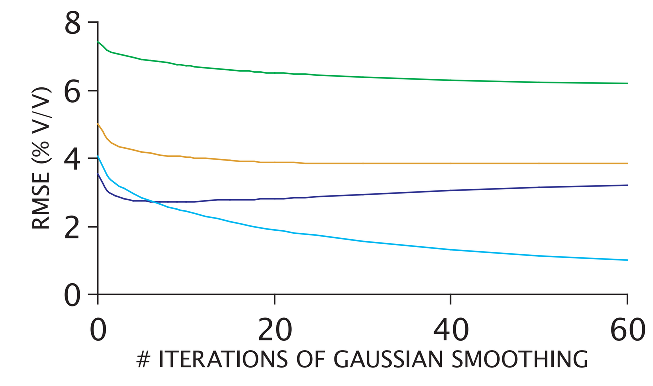
## MULTISCALE SIMILARITY

### UPSCALING BOTH FIELDS




### UPSCALING OBSERVATION




Common measures of similarity work only at the broadest scale (e.g. comparing means) or the finest scale (e.g. comparing every cell) of the field. However, the scales in between provide additional information that is useful for comparison. We do this comparison visually, but not quantitatively.

Multiscale measures plot similarity against scale, allowing the user to understand at which scales similarity is highest. If both fields have equivalent scales, then upscaling both fields shows if there is improved similarity at coarser scales.

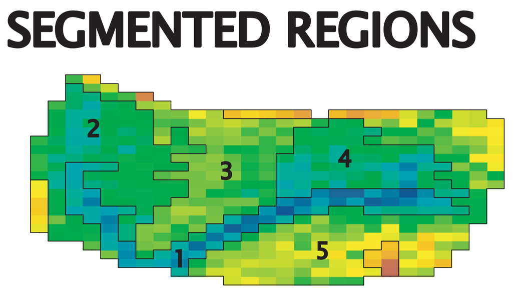
If the scales are different between two fields (e.g. point versus area averages), multiscale measures can be used to reveal the scale at which they agree. By only upscaling one field, the user can observe if there is a characteristic scale where the similarity improves.

NOTE THE DROP IN RMSE AS WE UPSCALE THE OBSERVATION.

THIS SUGGESTS INCONSISTENCY OF SCALES BETWEEN THE FIELDS AT THE FINEST RESOLUTION.

## WEIGHTED SIMILARITY

### SEGMENTED REGIONS




A SIMILARITY MEASURE IS CALCULATED FOR EACH REGION IN THE OBSERVED.

THIS HELPS THE USER ASSESS WHICH FEATURES ARE WELL SIMULATED.

| REGION | RMSE (% V/V) 1. | 2. | 3. |
|---|---|---|---|
| 1 | 2.7 | 2.8 | 12.2 |
| 2 | 2.5 | 3.0 | 8.1 |
| 3 | 4.6 | 6.3 | 4.0 |
| 4 | 2.9 | 2.8 | 9.1 |
| 5 | 3.2 | 5.5 | 4.7 |

### WEIGHTING FIELDS


USING A WEIGHTING FIELD, THE SIMILARITY MEASURE FOCUSSES ON SPECIFIC TOPOGRAPHIC FEATURES.

| FOCUS | RMSE (% V/V) 1. | 2. | 3. |
|---|---|---|---|
| North aspect | 2.3 | 3.4 | 4.8 |
| Steep slopes | 2.8 | 3.8 | 5.8 |

Most similarity measures treat every element equally during comparison. However, with visual comparison we focus on particular parts of the spatial field more than others. Similarity measures replicate this by applying higher weights to important elements, thus focussing the meaning of the measure.

Weights can be specified subjectively to make a similarity measure test a characteristic of the spatial model (e.g. weighting north facing slopes highly to assess performance of evaporation). Weights can also be applied to limit the calculation of the similarity measure. By doing so, the meaning of the similarity measure can be refined.

## SUMMARY

There are many avenues for improving spatial field comparison. Current methods are useful for comparing any data set, but for harnessing the spatial arrangement of values in a field, alternatives methods must be used.

This work shows: 1) how tolerances for differences in value and location can be considered during comparison; 2) how analysis based on regions can be achieved with spatial fields; 3) how weighting clarifies the meaning of a similarity measure; and 4) how comparison at multiple scales provides additional information to the user. All these methods utilise the spatial arrangement of values to produce quantitative measures that compare spatial fields. For further information, please contact the authors or refer to:


Wealands, S.R., Grayson, R.B. and Walker, J.P. (in press), “Quantitative comparison of spatial fields for hydrological model assessment – some promising approaches”, Advances in Water Resources.



**Stephen Wealands**

Department of Civil and Environmental Engineering,
The University of Melbourne, AUSTRALIA 3010

Email: srweal@civenv.unimelb.edu.au

COOPERATIVE RESEARCH CENTRE FOR CATCHMENT HYDROLOGY

THE UNIVERSITY OF MELBOURNE

**Author Affiliations**

a Department of Civil and Environmental Engineering, The University of Melbourne, Australia
b Cooperative Research Centre for Catchment Hydrology, Australia
c Inst. of Hydraulics, Hydrology and Water Resources Management, Vienna University of Technology, Austria