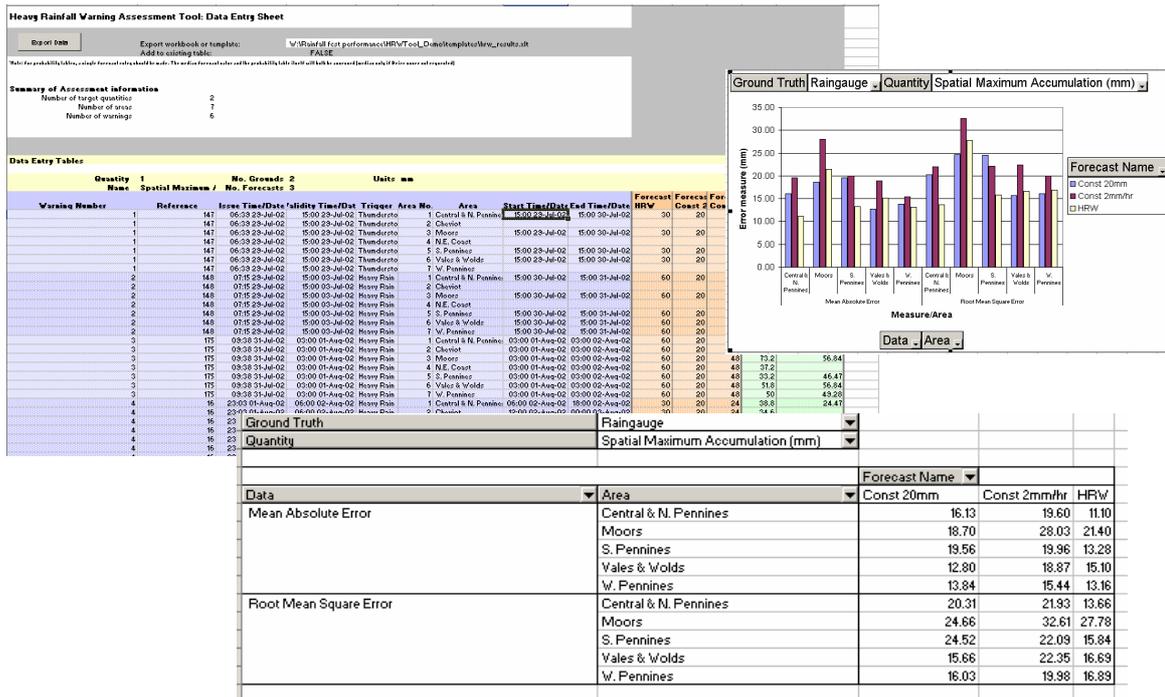


Heavy Rainfall Warning Assessment Tool



User Guide

Version 1.2



Commissioning Organisation

Environment Agency
Rivers House
Waterside Drive
Aztec West
Bristol BS32 4UD

Tel: +44(0)1454 624400

Fax: +44(0)1454 624409

Met Office
London Road
Bracknell
Berkshire
RG12 2SY

Tel: +44(0)1344 855680

Fax: +44(0)1344 855681

© Environment Agency, Met Office and CEH Wallingford

Jul 2004

All rights reserved. No part of this document may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the Environment Agency, the Met Office and CEH Wallingford.

The views expressed in this document are not necessarily those of the Environment Agency, the Met Office or CEH Wallingford. Their officers, servants or agents accept no liability whatsoever for any loss or damage arising from the interpretation or use of the information, or reliance upon views contained herein.

Dissemination status

Internal: Released Internally

External: Released to Public Domain.

Statement of Use

This User Guide to the Heavy Rainfall Warning Assessment Tool is an output of a Project that aimed to develop rainfall forecast performance monitoring criteria for use by the Met Office and the Environment Agency.

Contractor

This document was produced under Met Office Contract Number PB/B3734 by CEH Wallingford. The CEH Project Manager was R J Moore.

Authors: A E Jones, D A Jones, R J Moore and A J Robson

Project Leader

The Project Leader for Contract PB/B3734 was: Bryony May, Joint Centre for Hydro-Meteorological Research, Wallingford

The Environment Agency Client was: Alison Pickles, National Flood Warning Centre

Further copies of this guide are available from:
CEH Wallingford, Maclean Building, Crowmarsh Gifford,
Wallingford, Oxon, OX10 8BB, UK
Tel: 01491-838800 Fax: 01491-692424 e-mail: rm@ceh.ac.uk



**Centre for
Ecology & Hydrology**

NATURAL ENVIRONMENT RESEARCH COUNCIL

Contents

1	Overview of the Assessment Tool	1
1.1	System Requirements	1
1.2	Installation Procedure	1
1.3	Assessment Procedure	3
1.4	About this User Guide	3
1.5	Limitations	3
2	Configuration of a Heavy Rainfall Warnings Assessment	4
2.1	Creating a new Data Entry Workbook	4
2.2	Overview of the Configuration Sheet	4
2.3	Assessment Details	5
2.3.1	Reference	5
2.3.2	Notes	6
2.3.3	Start Date	6
2.3.4	End Date	6
2.3.5	Number of areas to be assessed	6
2.3.6	Names of areas	6
2.3.7	Number of target quantities	6
2.3.8	Names of target quantities	6
2.3.9	Units	7
2.3.10	Number of warnings to assess	7
2.4	Performance Measures	7
2.5	Target Quantities	7
2.5.1	Quantity	7
2.5.2	Name	7
2.5.3	Units	7
2.5.4	Number of forecasts to assess	7
2.5.5	Names of forecasts	7
2.5.6	Type TRUE if forecast is in probability form	8
2.5.7	Number of rainfall categories for probability tables	8
2.5.8	Number of ground-truths	8
2.5.9	Names of ground-truths	8
2.5.10	Number of skill score thresholds	8
2.5.11	Skill score threshold values	8
2.6	Creating a pre-configured template	8
3	Data Entry	9
3.1	Header information	10
3.2	Data Entry Tables	10
3.2.1	Warning Number	10
3.2.2	Reference	10
3.2.3	Issue Time/Date	11
3.2.4	Validity Time/Date	11
3.2.5	Trigger	11
3.2.6	Area No.	11
3.2.7	Area	11
3.2.8	Start Time/Date	11
3.2.9	End Time/Date	11
3.2.10	Forecast(s)	11
3.2.11	Probability Forecasts	11
3.2.12	Ground-truth(s)	12

3.2.13	Notes	12
3.3	Export Data Button	12
4	Export Data Table Creation	13
4.1	Creating a new export workbook	14
4.2	Adding data to an existing workbook	14
4.3	The Results Workbook	14
5	Results Tables and Charts	15
5.1	Generating Results	15
5.2	Results tables	16
5.3	Summary tables	16
5.4	Single warnings	21
5.5	Charts	23
5.5.1	Charts in Excel 97	23
5.5.2	Charts in Excel 2000/2002	27
6	Guide to Performance Measures	30
6.1	Guide to Assessment of Heavy Rainfall Warnings	31
6.1.1	Selecting Target Quantities for Assessment	31
6.1.2	Selecting and Determining Ground-truths for Assessment	32
6.1.3	Selecting Forecasts	32
6.1.4	Practicalities and limitations of using the Assessment Tool	33
6.2	Guide to Performance Measures, Part 1	34
6.2.1	Example	34
6.2.2	Guide to Notation	35
6.2.3	Continuous Performance Measures	35
6.2.4	R ² Efficiency	37
6.2.5	Categorical Skill Scores	37
6.3	Guide to Performance Measures, Part 2	40
6.3.1	Relative Categorical Skill Scores	40
6.3.2	Skill Scores for Probability Forecasts	42
6.4	Guide to Making Comparisons with the HRW Assessment Tool	43
6.4.1	Guide to Comparing Forecast Sources	43
6.4.2	Guide to Comparing Ground-truths	44
7	Example Assessments	45
7.1	Northeast Region	45
7.1.1	Heavy Rainfall Warnings	45
7.1.2	Ground-truth	48
7.2	Northwest Region	50
7.2.1	Heavy Rainfall Warnings	50
7.2.2	Ground-truth	52
7.3	Thames Region	53
7.3.1	Heavy Rainfall Warnings	53
7.3.2	Ground-truth	61
8	HRW Assessment Tool FAQ	62
	References	62
	Appendix A Probability interpretation of Relative Categorical Skill Scores	63
	Appendix B Glossary of terms	64

1 Overview of the Assessment Tool

The Heavy Rainfall Warning (HRW) Assessment Tool is a toolkit for Microsoft Excel. The tool allows the user to configure an assessment framework for a particular format of Heavy Rainfall Warning, enter and save data for forecasts and ground-truths, and generate a range of performance measures and other statistics for new and previously saved data. Summary tables are presented using Excel's PivotTable feature, from which charts can also be generated.

1.1 System Requirements

The HRWAssessment Tool runs under Microsoft Excel 97 or later on Windows 95/98/ME, Windows NT, Windows 2000 and Windows XP. There are slight differences between the workbooks generated for Excel 97, 2000 and 2002. For normal use the following minimum system specification is recommended:

- Desktop workstations
- At least 300 MHz Pentium
- At least 64 Mbytes physical memory
- At least 20 Mbytes free disk space
- A display capable of 800x600 by 256 colours
- A mouse or other compatible pointing device.

1.2 Installation Procedure

This version of the HRW Tool may be run entirely or partially from the CD.

It is recommended that the hrw_tool folder is copied from the CD and stored locally. The templates and examples supplied on the HRW Tools CD have been setup with the default of expecting the hrw_tool folder to reside under c:\hrw_assessment. If this folder is stored elsewhere then the templates and examples should be edited and saved so that they point to the correct locations. This is described below. There should be no need to edit the templates and examples if the c:\hrw_assessment directory is chosen.

If the templates are used directly from the CD the location of the workbook or template will have to be specified each time the tool is used.

The HRW tool is designed so that dll files do not need to be copied from the CD. The programs will look for the dll files on the D drive: thus users will need access to the CD whenever they run the HRW tool. If the dll files are located elsewhere, or the D drive is not the CD-drive, then the HRW_results template should be modified to point to the correct location (see below).

Editing the HRW.xlt template

This operation will be necessary unless the templates are located in the c:\hrw_assessment\hrw_tool\templates directory.

- (1) Make sure that the templates directory has been moved to the required location.
- (2) Set the HRW.xlt file to have read/write access (Right click on the templates folder in Windows Explorer and select properties: make sure that the read-only box is not checked).

- (3) Open the HRW.xlt template by *right-clicking on the file name in Windows Explorer and selecting "Open"*. Do not double click to open, as this will simply generate a spreadsheet derived from the template.
- (4) Specify the location of the HRW_RESULTS.xlt template at the top of the "Data Table" worksheet of the HRW template, in the field labelled "Export workbook or template", as shown in Figure 1. The Browse button can be used to navigate to the location of the Results template as installed on your machine.
- (5) Save the HRW.xlt file
- (6) Make the HRW.xlt directory is read-only - right click on the templates folder in Windows Explorer and select properties, check the read-only box .

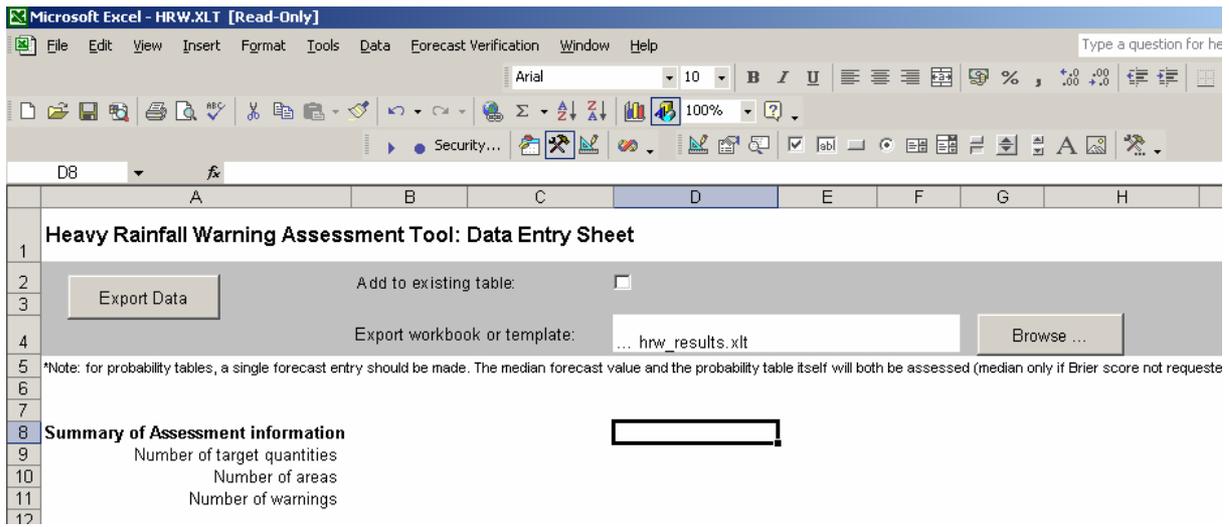


Figure 1 Specifying the location of the Results workbook template.

Editing the HRW_results.xlt template

This operation should only be needed if the dll is not to be found on the HRW tool CD under the D drive

- (1) –(3) as above, but for the HRW_results.xlt sheet.
- (4) The location of the forecasting verification dll should be specified at the top of the "Data" worksheet Results template, as shown in Figure 2. Again a Browse button is provided and should be used to navigate to the location of the forecast_verification.dll, located on the CD under \hrw_tool_dll\dll.
- (5)-(6) Save and make read-only as above.

Editing the examples workbooks

The examples will need to be edited unless the template files are stored in the c:\hrw_assessment\hrw_tool\templates directory. The examples should be edited in the same way as the HRW and HRW_results templates.

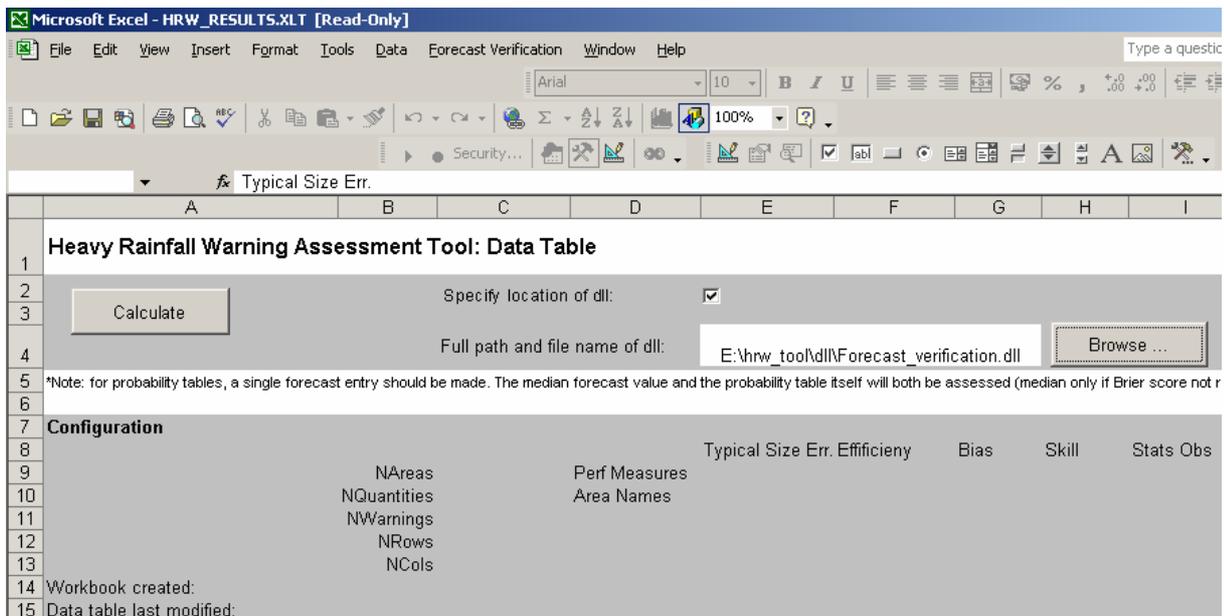


Figure 2 Specifying the location of the dll.

1.3 Assessment Procedure

Assessment using the HRW Tool consists of four stages: configuration, data entry, data table export and performance measure generation. Two different Excel workbooks are used in this process: the Data Entry Workbook for configuration and data entry, and the Results Workbook to which data are exported and in which results are calculated.

The starting point for assessment is the HRW.xlt Excel template. This is used to generate the Data Entry Workbook. The user carries out configuration and data entry in this workbook and then uses embedded worksheet buttons to export the data to the Results Workbook where performance measures and other statistics are generated. More information about the recommended procedure in terms of selection of ground-truths, comparative forecasts and sample sizes required for different performance measures is given in Section 6.

1.4 About this User Guide

Sections 2 to 5 of the User Guide present a "walk-through" of each aspect of the HRW Assessment Tool user interface, presented in the order typically encountered by the user. In Section 6 an overview of the recommended assessment procedure and a guide to each of the available performance measures is given. In Section 7 the data used in three example assessments are listed and advice is given as to how to make use of the information contained in the warnings. The Excel workbooks containing these example assessments can be found on the CD under hrw_tool\examples. A pdf version of this User Guide can be found on the CD under hrw_tool\documentation.

1.5 Limitations

It is important that the user does not alter the layout of the HRW Tool workbooks in any way. The correct functioning of the workbook is dependent on the worksheet layout being unchanged.

2 Configuration of a Heavy Rainfall Warnings Assessment

The first stage in the assessment process is configuration of the assessment. Sections 2.1 to 2.4 describe how to create a new Data Entry Workbook and enter configuration information. Many of the configuration specifications may remain constant for a given format of forecasts and so the construction of a pre-configured template is described in Section 2.6.

2.1 Creating a new Data Entry Workbook

To create a new Data Entry Workbook, navigate to the hrw_tool\templates folder and double click the file "hrw.xlt". Excel will open and create a new workbook based on the hrw.xlt template. This workbook is used to configure an assessment and enter data. You may wish to save the new workbook immediately with an appropriate name and location.

The new assessment workbook consists of two worksheets named "Configuration" and "Data Table". The Data Table Sheet is described in Section 3. Sections 2.2. to 2.5 describe how to fill in each part of the Configuration Sheet to configure an assessment.

2.2 Overview of the Configuration Sheet

An example Configuration Sheet filled in for a set of Northeast Region Heavy Rainfall Warnings is given in Figure 2.

The worksheet is headed with the title "Heavy Rainfall Warning Assessment Tool: Configuration Sheet". Below this title is a grey bar containing a button labelled "Create Tables". This button is used to create data entry tables on the "Data Table" sheet when the Configuration Sheet has been completed.

Below the grey bar is a purple area where the configuration details are filled in. Depending on the resolution of your monitor and font settings, you may wish to view this sheet at 75 % zoom in order to view the full width of the configuration area without scrolling.

The configuration area contains three main sections which defined the configuration. Sections 2.3 to 2.5 describe the three areas in turn.

When the configuration details have been entered, the Create Tables button should be clicked to generate the data entry tables on the Data Tables Sheet. See Section 3 for further information on data entry.

a) Top half - Assessment Details and Performance Measure Specification

	A	B	C	D	E	F	G	H
1	Heavy Rainfall Warning Assessment Tool: Configuration Sheet							
2	Create Tables							
3								
4								
5	*Note: for probability tables, a single forecast entry should be made. The median forecast value and the probability table itself will both be assessed (median only if Brier score not requested)							
6								
7	Assessment Details							
8	Reference	NE Event 5b Part 1						
9	Notes							
10	Start Date	29 July 2002						
11	End Date	31 July 2002						
12	Number of areas to be assessed	7						
13	Names of areas	Central & N. Pennines	Cheviot	Moors	N.E. Coast	S. Pennines	Vales & Wolds	W. Penn
14	Number of target quantities to be assessed	2						
15	Reference number of target quantities	1	2	3	4	5	6	7
16	Names of target quantities	Spatial Maximum Accumulation	Spatial Mean Accumulation					
17	Units	mm	mm					
18	Number of warnings to assess	3						
19	Performance Measures							
20	Typical size of error	TRUE						
21	Efficiency	TRUE						
22	Bias	TRUE						
23	Skill Scores	TRUE						
24	Statistics of observations	TRUE						
25	Statistics of forecasts	TRUE						
26	Compare forecasts	TRUE						
27	Compare ground truths	TRUE						
28	Brier Score (Probability)*	FALSE						
29								
30								

b) Bottom half - Target Quantities

31	Target Quantities							
32	Quantity	1						
33	Name	Spatial Maximum Accumulation						
34	Units	mm						
35	Number of forecasts to assess	3						
36	Names of forecasts	HRW	Const 20mm	const 2mm/hr				
37	Type TRUE if forecast is in probability table form							
38	Number of rainfall categories for probability tables							
39	Number of ground truths	2						
40	Names of ground truths	Raingauge	Radar					
41	Number of skill score thresholds	6						
42	Skill score threshold values >	0	14	29	39	49	59	
43								
44	Quantity	2						
45	Name	Spatial Mean Accumulation						
46	Units	mm						
47	Number of forecasts to assess	3						
48	Names of forecasts	HRW	Const 20mm	const 2mm/hr				
49	Type TRUE if forecast is in probability table form							
50	Number of rainfall categories for probability tables							
51	Number of ground truths	2						
52	Names of ground truths	Raingauge	Radar					
53	Number of skill score thresholds	6						
54	Skill score threshold values >	0	14	29	39	49	59	
55								

Figure 2 Example Configuration Sheet in the Data Entry Workbook.

2.3 Assessment Details

This section of the Configuration Sheet should be filled in with the relevant details for the assessment. Many of the fields entered here will be common to all assessments for a given format of Heavy Rainfall Warning. See Section 2.6 for more information about how to create a pre-configured template to avoid re-entering common information.

2.3.1 Reference

Optional. Enter a reference number, name or code for this assessment. This field will appear in the output and provides a way of identifying a set of assessment data. This field is not used for calculation by the software.

2.3.2 Notes

Optional. Enter any further information required about the assessment. This field is not used for calculation by the software.

2.3.3 Start Date

Optional. Enter the start date of the assessment period, if required. The date will be reformatted to the form "01 January 2002" and will appear in the output. This field is not used for calculation by the software.

2.3.4 End Date

Optional. Enter the end date of the assessment period, if required. The date will be reformatted to the form "01 January 2002" and will appear in the output. This field is not used for calculation by the software.

2.3.5 Number of areas to be assessed

Compulsory. Enter the number of separate areas for which Heavy Rainfall Warnings are issued and are to be assessed. When deriving performance measures, statistics will be calculated for each area. If desired, assessment can be carried out separately for each area, in which case "1" would be entered here. Note however that if this approach is taken there is no way to compare performance across different areas using the HRW Assessment Tool.

2.3.6 Names of areas

Compulsory. Enter a name for each area in a separate column. These area names appear in the output. Note area names and order of appearance in the list must remain consistent for different assessments if results are to be merged.

2.3.7 Number of target quantities

Compulsory. Enter the number of different target quantities to be assessed. A target quantity is a single quantity that is forecast in the Heavy Rainfall Warnings. This forecast quantity should always be able to be assessed against the same ground-truths and comparative forecasts. For example, in some regions the forecast quantity is always described as "rainfall amount", and so there would be just one forecast quantity. In other regions a rainfall amount and a maximum rainfall are usually given. Try to include all quantities that could be given in a warning, not just those included in the set of warnings to be assessed immediately.

2.3.8 Names of target quantities

Compulsory. Give a name for each of the target quantities. Here it is helpful to use a name that describes how the quantity would be derived from the true rainfall field. An example set of quantities is:

"Spatial Mean Accumulation" - the areal average rainfall accumulated over the appropriate time.

"Spatial Maximum Accumulation" - the areal maximum point rainfall in the accumulated field.

"Spatial Maximum Rainfall Rate" - the areal maximum point rainfall rate during the period.

Performance measures will be pooled across the same target quantity forecast in different warnings.

Note that for some regions (e.g. Thames/Southern/Anglian) the format of Heavy Rainfall Warnings includes two target quantities, each of which have two forecasts: a probability forecast and a 'most likely' forecast. See Section 7 and the accompanying example workbooks for further examples of target quantities contained in Heavy Rainfall Warnings from various regions.

Note quantity names and order of appearance in the list must remain consistent for difference assessments if results are to be merged, and so it is very important to define the quantities clearly before starting any assessment.

2.3.9 Units

Compulsory. Enter the units for the quantity, usually "mm" or "mm/hr".

2.3.10 Number of warnings to assess

Compulsory. Enter the total number of warnings to assess. This will usually be the physical number of Heavy Rainfall Warnings issued in the assessment period. However there may be cases when a single warning contains more than one forecast of a single quantity, for example a forecast of spatial mean rainfall for two different time periods. In this case the simplest solution is to include this as two warnings. See Section 7.2 for an example of this.

2.4 Performance Measures

In this section of the Configuration Sheet there is a drop down list for each set of performance measures; setting the value to TRUE will result in the appropriate performance measure being calculated, setting it to FALSE will mean the statistics are not produced. Note that the Brier Score should only be used when the warnings contain probability forecasts. See Section 6 for a guide to each set of performance measures.

2.5 Target Quantities

The last section of the Configuration Sheet contains a sub-section for each of the target quantities named in the Assessment Details section.

2.5.1 Quantity

Compulsory. This is the reference number of the target quantity given in the Assessment Details section. This number is automatically filled in.

2.5.2 Name

Compulsory. The name for this target quantity. This field is filled in automatically from the name give in the Assessment Details section.

2.5.3 Units

Compulsory. The units for this target quantity. This field is filled in automatically from the units give in the Assessment Details section.

2.5.4 Number of forecasts to assess

Compulsory. The number of different forecasts to be assessed for this quantity, including the Heavy Rainfall Warning forecasts and any comparative forecasts (e.g. Nimrod/Gandolf) or naive forecasts. In the example shown in Figure 2 the Heavy Rainfall Warning forecast and two naive forecasts have been included for both forecast quantities.

2.5.5 Names of forecasts

Compulsory. Names for each of the forecasts to be assessed. For the Heavy Rainfall Warnings this would normally be the name given in the warning (e.g. "Most likely point maximum"), or simply "HRW" if there is only ever one Heavy Rainfall Warning forecast type for this quantity.

2.5.6 Type TRUE if forecast is in probability form

Compulsory for probability forecasts. For other forecasts this field can be left blank. Note only one probability forecast is allowed for each target quantity.

2.5.7 Number of rainfall categories for probability tables

Compulsory for probability forecasts. The number of rows in the probability table, including a row for a lower bound of zero, even if this is not included in the probability table.

2.5.8 Number of ground-truths

Compulsory. The number of ground-truths to be used in the assessment of this quantity. This may include truth derived from different sources, e.g. raingauge and radar, and different methods of obtaining the ground-truth value, e.g. median or mean raingauge.

2.5.9 Names of ground-truths

Compulsory. A name for each form of ground-truth. It is helpful for the name to be as descriptive as possible in defining how the ground-truth was derived.

2.5.10 Number of skill score thresholds

Compulsory if skill scores requested in Performance Measures section.

2.5.11 Skill score threshold values

Compulsory if skill scores requested in Performance Measures section. The values of the thresholds to be used. These are used to derive measures of the performance of forecasts in predicting rainfall above or below a given amount or rate. A range of thresholds should be specified, including zero and the Heavy Rainfall Warning trigger values if known. Note that the skill scores look for forecasts and observation greater than the threshold value, so a common forecast of 30 mm is best tested using a threshold slightly less than 30 mm, e.g. 29, 29.5 or 29.9 mm. Units for the thresholds are the same as those of the target quantity.

2.6 Creating a pre-configured template

Many of the fields in the Configuration Sheet can be defined once for a given format of Heavy Rainfall Warning. This is advisable to ensure consistency across different assessments and to enable the merging of data to derive pooled measures over larger sample sizes. Once the fields have been decided upon, a template can be created as follows.

Open the HRW.xlt template file by right clicking on it in Windows Explorer and selecting Open. Now save the file under a different name as a template file (e.g. HRW_Thame.xlt). You may need to navigate back to the HRW Tool directory as Excel often has a pre-configured directory in which to save template files. Add the configuration details to the Configuration Sheet of the new template and save again as a template. Close the template and then open a workbook from it by double clicking on the file name in Windows Explorer. Check the configuration by clicking the Create Tables button. Make changes to the template file as necessary. The new template can now be used in place of HRW.xlt to generate new Data Entry Workbooks.

3 Data Entry

When the configuration has been completed, click the Create Tables button on the Configuration Sheet. This triggers a process to validate the configuration information and then use the information provided to create a set of data entry tables on the Data Table Sheet. Note that once the data tables have been generated, clicking the button again will overwrite the tables according to the current configuration. This is not recommended. To amend the data table structure, generate a new Data Entry Workbook from the HRW template and then copy and paste the relevant fields into the Configuration Sheet of the new workbook. Then use the Create Tables button to generate new data tables.

An example Data Table Sheet is shown in Figure 3. Sections 3.1 to 3.3 describe how to add data to the tables.

a) Data Table Sheet derived from the configuration shown in Figure 2.

Quantity	Name	Spatial Maximum Acc	No. Grounds	Units	Forecast 1 HRW	Forecast 2 Coast 20mm	Forecast 3 Coast 2mm/hr	Ground truth Rain gauge	Ground truth Radar				
1	147	06:33 23-Jul-02	15:00 23-Jul-02	Thundersto	1	Central & N. Pennine	15:00 23-Jul-02	15:00 30-Jul-02	30	20	48	20.6	58.03
1	147	06:33 23-Jul-02	15:00 23-Jul-02	Thundersto	2	Cheviot	15:00 23-Jul-02	15:00 30-Jul-02	30	20	48	2.2	34.63
1	147	06:33 23-Jul-02	15:00 23-Jul-02	Thundersto	3	Moors	15:00 23-Jul-02	15:00 30-Jul-02	30	20	48	35.8	183.00
1	147	06:33 23-Jul-02	15:00 23-Jul-02	Thundersto	4	N.E. Coast	15:00 23-Jul-02	15:00 30-Jul-02	30	20	48	16.2	136.66
1	147	06:33 23-Jul-02	15:00 23-Jul-02	Thundersto	7	W. Pennines	15:00 23-Jul-02	15:00 30-Jul-02	30	20	48	60.4	85.59
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	1	Central & N. Pennine	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	60.4	85.59
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	2	Cheviot	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	33.8	83.25
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	3	Moors	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	3.2	63.84
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	4	N.E. Coast	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	68.8	102.78
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	6	Vales & Wolds	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	33.8	83.25
2	148	01:15 23-Jul-02	15:00 03-Jul-02	Heavy Rain	7	W. Pennines	15:00 30-Jul-02	15:00 31-Jul-02	60	20	48	30	62.38
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	1	Central & N. Pennine	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	37	54.38
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	2	Cheviot	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	3.6	56.84
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	3	Moors	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	37.2	46.47
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	4	N.E. Coast	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	51.8	56.84
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	5	S. Pennines	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	30	43.28
3	175	03:38 31-Jul-02	03:00 01-Aug-02	Heavy Rain	6	Vales & Wolds	03:00 01-Aug-02	03:00 02-Aug-02	60	20	48	38.8	24.47
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	1	Central & N. Pennine	06:00 02-Aug-02	18:00 02-Aug-02	30	20	24	34.6	17.03
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	2	Cheviot	06:00 02-Aug-02	06:00 03-Aug-02	30	20	24	24.4	17
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	3	Moors	06:00 02-Aug-02	18:00 02-Aug-02	30	20	24	13.8	
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	4	N.E. Coast	18:00 02-Aug-02	00:00 03-Aug-02	30	20	24		
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	5	S. Pennines	06:00 02-Aug-02	18:00 02-Aug-02	30	20	24	12.6	17.03
4	16	23:03 01-Aug-02	06:00 02-Aug-02	Heavy Rain	6	Vales & Wolds	06:00 02-Aug-02	18:00 02-Aug-02	30	20	24	23.2	25.38
5	47	08:49 04-Aug-02	08:49 04-Aug-02	Heavy Rain	1	Central & N. Pennine	08:00 04-Aug-02	20:00 04-Aug-02	15	20	24	10	15.34
5	47	08:49 04-Aug-02	08:49 04-Aug-02	Heavy Rain	2	Cheviot	08:00 04-Aug-02	20:00 04-Aug-02	15	20	24	10	13.63
5	47	08:49 04-Aug-02	08:49 04-Aug-02	Heavy Rain	3	Moors	08:00 04-Aug-02	20:00 04-Aug-02	15	20	24	10	13.63
5	47	08:49 04-Aug-02	08:49 04-Aug-02	Heavy Rain	4	N.E. Coast	08:00 04-Aug-02	20:00 04-Aug-02	15	20	24	10	13.63
5	47	08:49 04-Aug-02	08:49 04-Aug-02	Heavy Rain	5	S. Pennines	08:00 04-Aug-02	20:00 04-Aug-02	15	20	24	10	13.63

b) Section of Data Table Sheet showing columns for entry of probability forecast data, example for Thames region "Probability of at least this Rainfall Rate" forecast.

Forecast	Probability of rainfall rate (Prob.)	Lower Bound 1	Probability (%)	Lower Bound 2	Probability (%)	Lower Bound 3	Probability (%)	Lower Bound 4	Probability (%)	Lower Bound 5	Probability (%)	Median	Ground t raingaug
15	0	30	4	90	10	60	20	40	50	50	10	15	8
18	0	30	4	90	10	60	20	40	50	50	10	15	7
12	0	80	4	80	10	40	20	10	50	50	0	8.5	10
8	0	100	4	90	10	30	20	10	50	50	0	8	20
8	0	100	4	90	10	30	20	10	50	50	0	8	5
32	0	100	4	100	10	100	20	70	50	10	30	25	
32	0	100	4	100	10	100	20	70	50	10	30	45	

Figure 3 Example Data Table Sheets.

3.1 Header information

The Data Table Sheet is headed with the title "Heavy Rainfall Warning Assessment Tool: Data Entry Sheet". Below this is a grey bar with a button labelled "Export Data" and two fields specifying how to export the data. This area is used when the data have been entered in the table. The path specified in the field labelled "Export workbook or template" can be either a template for a new workbook to export the data to, or an existing export workbook containing previously loaded data. The field marked "Add to existing table" specifies whether the export should be to a new workbook using the template or to an existing workbook. The export process is described fully in Section 4.

Below the grey bar is a white area containing a summary of the assessment details used to construct the data entry tables. This area is for information only; the configuration sheet is used when the data are exported.

3.2 Data Entry Tables

A yellow bar marks the start of the data entry tables on the sheet. Each target quantity specified in the configuration has a separate table headed by a yellow bar containing the number, name, units, number of ground-truths and number of forecasts for this target quantity. The header details are for information only; the configuration sheet is used when the data are exported.

As shown in Figure 3(a), each data entry table has a set of rows for each warning specified in the configuration. Within each warning section there are rows for each area. A single row is therefore used to represent a single target quantity for a single warning and single area. In each row there are columns to enter information about the Heavy Rainfall Warning (shaded purple), values for each single-valued forecast and probability forecast (shaded orange and blue respectively) and values for each form of ground-truth (shaded green).

Because rows in the table are generated for every combination of quantity, warning number and area, there may be a number of rows in the table that do not correspond to an existing warning. In this case the cells should simply be left blank. Blank cells are ignored when the data are exported for assessment. Similarly cells corresponding to missing data should be left blank.

In the current version of the HRW Assessment Tool, a row in the data entry table is ignored in the calculation of performance measures if any of the forecasts or ground-truths values are missing. This will not be the case in the release version of the software.

Sections 3.2.1 to 3.2.13 describe each column in the data entry table. Note that when entering data in the table, common fields will be copied automatically from the first row in the table or table section to all other relevant rows for convenience.

3.2.1 Warning Number

Compulsory. This gives the number of the warning within the current assessment and is filled in automatically. Do not edit this field.

3.2.2 Reference

Optional. This field can be used to store the warning reference number or code, if applicable. This field is not used for calculation by the software.

3.2.3 Issue Time/Date

Optional. This field can be used to store the issue time and date of the warning, if applicable. Both a time in the 24 hour clock and a date with year should be entered. This field is not used for calculation by the software.

3.2.4 Validity Time/Date

Optional. In some regions, e.g. North West, warnings are given a “valid time”: this may be the time up until which the warning is valid, or a period of validity. In regions where a validity time is stated, the validity field may be used to store the validity time and date of the warning. Both a time in the 24 hour clock and a date with year should be entered. The field is only provided for reference purposes and is not used for calculation by the software.

3.2.5 Trigger

Optional. This field can be used to store the trigger for warning, if applicable. This might be the type of warning e.g. Thunderstorm, or Heavy Rainfall, or might refer a specific type of event and use local terminology, e.g. "Event A" in Northwest region. This field is not used for calculation by the software.

3.2.6 Area No.

Compulsory. This field is automatically filled with the reference number of the area. Do not edit this field.

3.2.7 Area

Compulsory. This field is automatically filled with the name of the area. Do not edit this field.

3.2.8 Start Time/Date

Optional. This field can be used to store the start time and date of the warning, to aid in the manual data entry of ground-truth and forecast values. For ease of reference the time is best converted to GMT before entry. Both a time in the 24 hour clock and a date with year should be entered. This field is not used for calculation by the software.

3.2.9 End Time/Date

Optional. This field can be used to store the end time and date of the warning, to aid in the manual data entry of ground-truth and forecast values. For ease of reference the time is best converted to GMT before entry. Both a time in the 24 hour clock and a date with year should be entered. This field is not used for calculation by the software.

3.2.10 Forecast(s)

Each single-valued forecast (i.e. not probability forecast) specified in the configuration has a column headed by the configured forecast name. The forecast value for the corresponding quantity and area should be entered here. Missing or non-existent values should be left blank. Units are those given in the header bar for the target quantity.

3.2.11 Probability Forecasts

As shown in Figure 3(b), if a probability forecast is specified in the configuration, columns shaded blue appear in the data entry table. From left to right in the data entry table, the probability table should be entered in the form: lower bound 1, forecast 1 (%), lower bound 2, forecast 2 (%), etc. The lower bounds must increase from left to right and the probabilities must stay constant or decrease. The first lower bound must always be equal to zero, even if this does not appear in the Heavy Rainfall Warning probability table. For the current format of Thames/Southern/Anglian forecasts, this means that a probability for a rainfall rate greater than zero needs to be inferred from

the value for a lower bound of zero in the accompanying "Probability of Rainfall Amounts" table. For an example, see Section 7 and the accompanying example workbooks.

Any probabilities left blank in the table are treated as zero. The estimated median (50% point) of the entered distribution is calculated automatically and appears in a column to the right of the probability table values. If the median value does not appear, check the spreadsheet cell values have been recalculated by pressing the F9 key. If still no median appears, the entered probability table may be invalid.

3.2.12 Ground-truth(s)

Each form of ground-truth specified in the configuration has a corresponding column in the table to enter the ground-truth value. The units are the same as those specified in the header for the target quantity. For more information on the computation of ground-truth to enter into the table, see Section 6.

3.2.13 Notes

The final column in the table is headed "Notes" and provides space for any comments on the values entered. This field is not used for calculation by the software.

3.3 Export Data Button

When the data entry has been completed, the next step is to export the data to a new workbook where calculation of performance measures and other statistics is carried out. The export process is carried out automatically by the tool when the "Export Data" button is clicked. The export process is described in Section 4.

4 Export Data Table Creation

The third stage in the assessment process is the export of the data from the Data Table Sheet of the Data Entry workbook to a Results Workbook where performance measures are calculated. The export process is carried out automatically by the tool when the "Export Data" button on the Data Table Sheet is clicked.

There are two ways to export the data: to a new Results Workbook or to an existing Result Workbook that has been previously created using the HRW Assessment Tool. The second method is used when new data need to be combined with previously loaded data in order to derive performance measures from the combined dataset. When adding to an existing workbook both datasets must have the same configured areas, quantities, forecasts and ground-truths. Some consistency checking is performed by the software when the export to an existing workbook is requested.

Figure 4 shows an example Data Sheet in a Results Workbook.

Heavy Rainfall Warning Assessment Tool: Data Table

Calculate Specify location of dll: Full path and file name of dll: Browse ...

*Note: for probability tables, a single forecast entry should be made. The median forecast value and the probability table itself will both be assessed (median only if Brier score not requested)

Configuration

NAreas	7	Perf Measures	TRUE	Efficiency	TRUE	Bias	TRUE	Skill	TRUE	Stats Obs	TRUE	Stats Fcasts	TRUE	Cmp F	TRUE	Cmp D	TRUE
NQuantities	2	Area Names	Central & N. Penn	Cheviot	Moors	N.E. Coast	S. Pennines	Vales & Wolds	W. Pennines								
NWarnings	3																
NRows	132																
NCols	39																

Workbook created: #####
Data table last modified: #####

Data Table

Assessment Reference	Start Date	End Date	No. Areas	No. Quantities	Quantity	Quantity	Units	NumSSCats	SSCat1	SSCat2	SSC
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		
NE Event 5b Part 1	7/29/2002	7/31/2002	7	2	1	Spatial Ma mm	6	0	14		

Figure 4 Example Data Sheet created when data are exported

4.1 Creating a new export workbook

Below the title of the Data Table Sheet in the Data Entry Workbook is a grey bar containing the two fields which specify how to export the data. To export the data to a new Results Workbook, the path specified in the field labelled "Export workbook or template" should give the full pathname of the "hrw_results.xlt" export workbook template as installed on your system. This pathname should appear by default if the HRW Assessment Tool has been set up correctly. The check box labelled "Add to existing table" should be unchecked.

4.2 Adding data to an existing workbook

Below the title of the Data Table Sheet is a grey bar containing the two fields which specify how to export the data. To export the data to an existing Results Workbook, the path specified in the field labelled "Export workbook or template" should give the full pathname of the workbook to be used. The check box labelled "Add to existing table" should be checked.

Data can only be exported to a Results Workbook previously created using the HRW Assessment Tool. The datasets in the existing workbook and the data added must have the same configured areas, quantities, forecasts and ground-truths. Some consistency checking is performed when the export to an existing workbook is requested.

4.3 The Results Workbook

When the Export Data button is clicked, a new or existing Results Workbook is opened and the data from the Data Table Sheet in the Data Entry Workbook are written to the end of the large data table on the Data Sheet of the Results Workbook. An example is shown in Figure 4. Users of the HRW Assessment Tool will not normally need to look in detail at the contents of the table.

At the top of the Data Sheet in the Results Workbook is a button labelled "Calculate" which is used to calculate performance measures and other statistics from the contents of the data sheet. The configuration specified in the Data Entry Workbook is automatically written to the Results Workbook when the data are exported and so the Results Workbook can be used independently when calculating performance measures.

In order to store data for future use, the export workbook should be saved at this point, before the performance measures and other statistics are calculated. New data cannot be added to a workbook which contains calculated performance measures. If this form of the workbook is not saved, datasets should be merged in future by creating the Results Workbook again from the original Data Entry workbook, and then other datasets exported to the newly created Results Workbook.

5 Results Tables and Charts

The fourth and final stage in the assessment process is the calculation of performance measures and the generation of summary tables and charts.

5.1 Generating Results

Performance measures and other statistics can be calculated in any Results Workbook that contains a Data Sheet only. This workbook may have been newly generated from a set of data, recently added to with new data, or previously saved without results being calculated. Results are calculated based on solely the information and data contained in the Data Sheet and so the Results Workbook is completely self-contained.

Results are generated by clicking the "Calculate" button at the top of the Data Sheet in the Results Workbook. Performance measures and other statistics are automatically calculated and pairs of worksheets are added to the workbook, one pair for each of the groups of statistics requested in the configuration. Within such a pair of worksheets, one sheet contains a list-formatted table of results. The other sheet contains a summary table generated from the list-formatted table. The summary tables are presented using Excel's built in Pivot Table feature. Table 5 shows the names of the worksheets and the statistics contained in each.

For target quantites where there is only a single warning it is only possible to provide a restricted set of statistics. Section 5.4 details what is available for single warnings.

Table 5 Statistics generated using the HRW Assessment Tool

Configuration Name	Description	Results Table Sheet	Summary Sheet
Typical size of error	Mean absolute error and root mean square error.	Performance Statistics	PerformanceSummary
Efficiency	R ² statistic.		
Bias	Mean, median and percentage errors.		
Skill Scores	Categorical performance measures	Skill Scores	SkillScoresSummary
Statistics of observations	Mean, median and standard deviation of observations.	Stats Obs	StatsObsSummary

Statistics of forecasts	Mean, median and standard deviation of forecasts.	Stats Fcsts	StatsFcstsSummary
Compare forecasts	Standardised differences of Typical size of error measures across different forecasts.	Cmp Forecasts	CmpFcstsSummary
Compare ground-truths	Standardised differences of Typical size of error measures across different ground-truths.	Cmp G truths	CmpGSummary
Brier Score (Probability)	Continuous Brier Score for probability forecasts.	Brier Score	BrierSummary

5.2 Results tables

The results tables provide access to all the statistics generated by the HRW Assessment Tool and are used to generate the pivot tables presented in the summary sheets.

5.3 Summary tables

The summary sheets provide pivot tables of the results allowing the user to view the statistics generated in an easily readable form. The exact appearance of the pivot tables depends on the version of Excel used.

Users are advised to read the information about pivot tables in the Excel online help if they are not already familiar with this feature of the software. It is also advisable to keep a back-up copy of the automatically generated workbook before attempting to modify the pivot tables. Examples of the pivot tables generated for each set of statistics are given in Figures 5.3.1 to 5.3.7. These examples demonstrate the format of the pivot tables. More information about the interpretation of the statistics found in the tables is given in Section 6. A blank in a pivot table indicates that it was not possible to calculate this particular quantity.

By default, the page fields (those appearing at the top of the pivot table) in the pivot tables include "all" as one of the categories. Before using the values in the table, it is important to specify a single option (e.g. a ground-truth, a single quantity). This is to ensure that the values in the pivot tables correspond to unique values in the source tables. By default, Excel uses a summary function on the data, in this case the average. So for example, if "(all)" was specified in the Quantities field shown in Figure 5.3.1 instead of a single named quantity, the values in the table would be the average of the values for all quantities, which would be meaningless. The automatically generated pivot tables have been set up so that as long as a single option is specified for each of the page fields, the values in the tables should correspond to unique values in the source table.

Page fields - must be specified

Ground Truth	Raingauge				
Quantity	Spatial Maximum Accumulation (mm)				
		Forecast Name			
Data	Area	Const 20mm	Const 2mm/hr	HRW	
Mean Absolute Error	Central & N. Pennines	16.13	19.60	11.10	
	Moors	18.70	28.03	21.40	
	S. Pennines	19.56	19.96	13.28	
	Vales & Wolds	12.80	18.87	15.10	
	W. Pennines	13.84	15.44	13.16	
Root Mean Square Error	Central & N. Pennines	20.31	21.93	13.66	
	Moors	24.66	32.61	27.78	
	S. Pennines	24.52	22.09	15.84	
	Vales & Wolds	15.66	22.35	16.69	
	W. Pennines	16.03	19.98	16.89	
Mean Error (Bias)	Central & N. Pennines	9.47	-10.53	-8.03	
	Moors	0.50	-19.50	-17.00	
	S. Pennines	11.56	-11.64	-7.44	
	Vales & Wolds	2.40	-17.60	-15.10	
	W. Pennines	5.84	-12.56	-13.16	
% Error in Maximum Obs (%)	Central & N. Pennines	66.89	20.53	0.66	
	Moors	72.68	34.43	18.03	
	S. Pennines	70.93	30.23	12.79	
	Vales & Wolds	61.39	7.34	-15.83	
	W. Pennines	60.00	4.00	-20.00	
Median Error	Central & N. Pennines	8.80	-12.50	-7.20	
	Moors	-10.00	-26.00	-12.80	
	S. Pennines	13.20	-14.00	-5.00	
	Vales & Wolds	-5.60	-14.10	-15.60	
	W. Pennines	9.20	-14.00	-10.00	
R-squared for Root Mean Square Error (dimensionless)	Central & N. Pennines	-0.28	-0.49	0.42	
	Moors	0.00	-0.75	-0.27	
	S. Pennines	-0.29	-0.04	0.46	
	Vales & Wolds	-0.02	-1.09	-0.16	
	W. Pennines	-0.15	-0.79	-0.28	

Performance measures

Areas

Values of performance measures for each area and forecast

Figure 5.3.1 Pivot table generated for "Performance Statistics". Note units of performance measures are the same as that of the quantity unless specified otherwise in the name of the performance measure.

Quantity	Spatial Maximum Accumulation (mm)			
		Forecast		
Data	Area	Const 20mm	Const 2mm/hr	HRW
Median Forecast	Central & N. Pennines	20.00	48.00	30.00
	Moors	20.00	48.00	30.00
	S. Pennines	20.00	48.00	30.00
	Vales & Wolds	20.00	48.00	30.00
	W. Pennines	20.00	48.00	30.00
Mean Forecast	Central & N. Pennines	20.00	40.00	37.50
	Moors	20.00	40.00	37.50
	S. Pennines	20.00	43.20	39.00
	Vales & Wolds	20.00	40.00	37.50
	W. Pennines	20.00	38.40	39.00
St Dev Forecast	Central & N. Pennines	0.00	12.39	18.37
	Moors	0.00	12.39	18.37
	S. Pennines	0.00	10.73	20.12
	Vales & Wolds	0.00	12.39	18.37
	W. Pennines	0.00	13.15	20.12
NOccasions	Central & N. Pennines	6	6	6
	Moors	6	6	6
	S. Pennines	5	5	5
	Vales & Wolds	6	6	6
	W. Pennines	5	5	5

Figure 5.3.3 Pivot table for statistics of forecasts. The NOccasions field gives the sample size used to compute the performance measures.

A	B	C	D
Quantity	Spatial Maximum Accumulation (mm)		
		Ground Truth	
Data	Area	Radar	Raingauge
Median Observation	Central & N. Pennines	47.10	28.80
	Moors	25.84	10.00
	S. Pennines	51.88	33.20
	Vales & Wolds	41.42	14.40
	W. Pennines	48.66	29.20
Mean Observation	Central & N. Pennines	46.27	29.47
	Moors	33.43	20.50
	S. Pennines	85.02	31.56
	Vales & Wolds	66.25	22.40
	W. Pennines	39.93	25.84
St Dev of Observations	Central & N. Pennines	25.39	19.68
	Moors	22.32	27.01
	S. Pennines	64.21	24.17
	Vales & Wolds	68.97	16.95
	W. Pennines	19.86	16.69
NOccasions	Central & N. Pennines	6	6
	Moors	6	6
	S. Pennines	5	5
	Vales & Wolds	6	6
	W. Pennines	5	5

Figure 5.3.4 Pivot table for statistics of observations. The NOccasions field gives the sample size used to compute the performance measures.

Forecast	HRW			
Ground Truth	Radar			
Quantity	Spatial Maximum Accumulation (mm)			
		Base Forecast		
Data	Area	Const 20mm	Const 2mm/hr	
Mean Absolute Error	Central & N. Pennines	-2.39	0.14	
	Moors	-1.31	-2.42	
	S. Pennines	-1.86	1.63	
	Vales & Wolds	-1.54	-0.48	
	W. Pennines	-2.25	0.28	
Root Mean Square Error	Central & N. Pennines	-1.70	-0.09	
	Moors	-1.43	-1.57	
	S. Pennines	-1.85	0.91	
	Vales & Wolds	-1.93	0.76	
	W. Pennines	-1.90	0.32	

Figure 5.3.5 Pivot table for comparison of forecasts. This table shows the results of comparing each of the two naive forecasts with the Heavy Rainfall Warning forecast. A value larger than 2.2 in the table indicates strong evidence that the "Base Forecast" is better than the HRW Forecast. A value less than -2.2 indicates strong evidence that the "Base Forecast" is worse than the HRW forecast.

Ground Truth	Radar		
Forecast	HRW		
Quantity	Spatial Maximum Accumulation		
		Base Ground Truth	
Data	Area	Raingauge	
Mean Absolute Error	Lune	-0.80	
	South Lakes (1)	1.80	
	South Lakes (2)	-0.54	
	Upper Eden	-4.11	
	West Lakes	1.37	
Root Mean Square Error	Lune	-0.82	
	South Lakes (1)	1.40	
	South Lakes (2)	-0.49	
	Upper Eden	-2.75	
	West Lakes	0.98	

Figure 5.3.6 Pivot table for comparison of ground-truths. This table shows the results of comparing the use of each of the two sources of ground-truth to assess the performance of the Heavy Rainfall Warning forecast. A value larger than 2.2 in the table indicates strong evidence that the performance of the HRW forecast as measured using the raingauge ground-truth is better than the performance measured using the radar ground-truth. A value less than -2.2 indicates strong evidence that the opposite is true.

Ground Truth	raingauge			
Quantity	Spatial Maximum			
		Distribution	Forecast Name	
		100 % Error		Operational
Data	Area	Most likely point maximum	Probability of rainfall amount (median)	Probability of rainfall amount
Mean Absolute Error (certain)	Northeast	11.67	15.83	
	Southeast	10.00	12.67	
Continuous Brier Score	Northeast	6.99	11.41	9.72
	Southeast	7.32	8.54	8.81

Figure 5.3.7 Pivot table for Brier Score. The table shows the mean absolute error for the two single-valued forecasts (the Most likely point maximum and the median of the probability forecast), along with the continuous Brier Score for the single-valued forecasts with a 100% error band and the probability forecast itself.

5.4 Single warnings

Single warnings necessarily contain only a very limited amount of data and are of limited use in assessing heavy rainfall warnings. The Heavy Rainfall Warning Assessment Tool is primarily designed for use with multiple forecasts.

For a single warning only a limited number of statistics can be calculated. The most useful statistic (see section 6.2) is the bias (mean error). This is simply the difference between the observation and forecast. Note that the root mean square error and median error also give an identical value, as does the mean absolute error but without the sign. The R^2 Efficiency is undefined. The % error in the maximum observation provides a scaled, dimensionless measure of the size of error and thus is also of possible value.

Other statistics such as the mean of the forecasts or the mean of the observations are determined for single observations – but they simply reproduce the single value(s) from the warning. Skill Scores are not calculated for single observations – their values are undefined or are equal to 1 or 0 and are not useful. Forecast and ground-truth comparison statistics are also not provided as they are undefined for single observations.

Figures 5.4.1-5.4.4 show results that are obtained from the Tool when there are only data for a single warning. The data are based on a cut down version of the data illustrated in 7.2. For this instance, the Lune has only one warning, whilst the other 2 areas have data for 3 warnings. The Skill Scores summary pivot table is not included in the figures because it is blank for the Lune.

Ground Truth	Radar		
Quantity	Spatial Maximum Accumulation (mm)		
		Forecast Name	
Data	Area	Const 20mm	HRW
Mean Absolute Error	Lune	13.60	6.40
	Upper Eden	38.80	18.80
	West Lakes	28.33	8.33
Root Mean Square Error	Lune	13.60	6.40
	Upper Eden	40.00	19.17
	West Lakes	28.48	10.59
Mean Error (Bias)	Lune	13.60	-6.40
	Upper Eden	38.80	18.80
	West Lakes	28.33	8.33
% Error in Maximum Obs (%)	Lune	40.48	-19.05
	Upper Eden	70.24	25.60
	West Lakes	61.69	4.21
Median Error	Lune	13.60	-6.40
	Upper Eden	44.00	17.20
	West Lakes	27.40	5.40
R-squared (dimensionless)	Lune		
	Upper Eden	-15.98	-2.90
	West Lakes	-98.59	-12.79

Figure 5.4.1 Pivot table generated for "Performance Statistics". The Lune has only a single warning and thus the R-squared cannot be calculated.

Quantity	Spatial Maximum Accumulation (mm)		
		Forecast	
Data	Area	Const 20mm	HRW
Median Forecast	Lune	20.00	40.00
	Upper Eden	20.00	40.00
	West Lakes	20.00	40.00
Mean Forecast	Lune	20.00	40.00
	Upper Eden	20.00	40.00
	West Lakes	20.00	40.00
St Dev Forecast	Lune		
	Upper Eden	0.00	10.00
	West Lakes	0.00	10.00
NOccasions	Lune	1	1
	Upper Eden	3	3
	West Lakes	3	3

Figure 5.4.2 Pivot table for statistics of forecasts. The Lune has only one warning so for this site this table simply reports the observed values. The standard deviation cannot be calculated for the Lune, so this cell is blank.

Quantity	Spatial Maximum Accumulation (mm)		
		Ground Truth	
Data	Area	Radar	Raingauge
Median Observation	Lune	33.60	15.00
	Upper Eden	64.00	22.00
	West Lakes	47.40	22.00
Mean Observation	Lune	33.60	15.00
	Upper Eden	58.80	20.33
	West Lakes	48.33	20.33
St Dev of Observations	Lune		
	Upper Eden	11.89	4.73
	West Lakes	3.49	4.73
NOccasions	Lune	1	1
	Upper Eden	3	3
	West Lakes	3	3

Figure 5.4.3 Pivot table for statistics of observations. The Lune has only one warning so for this site this table simply reports the observed values.

Forecast	Const 20mm	
Ground Truth	Radar	
Quantity	Spatial Maximum Accumulation (mm)	
		Base Forecast
Data	Area	HRW
Mean Absolute Error	Lune	
	Upper Eden	3.46
	West Lakes	3.46
Root Mean Square Error	Lune	
	Upper Eden	2.76
	West Lakes	4.03

Figure 5.4.4 Pivot table for comparison of forecasts. This table shows the result of comparing the naive forecast with the Heavy Rainfall Warning forecast. The entry for the Lune is blank as it is not possible to calculate the comparison statistics with only 1 warning.

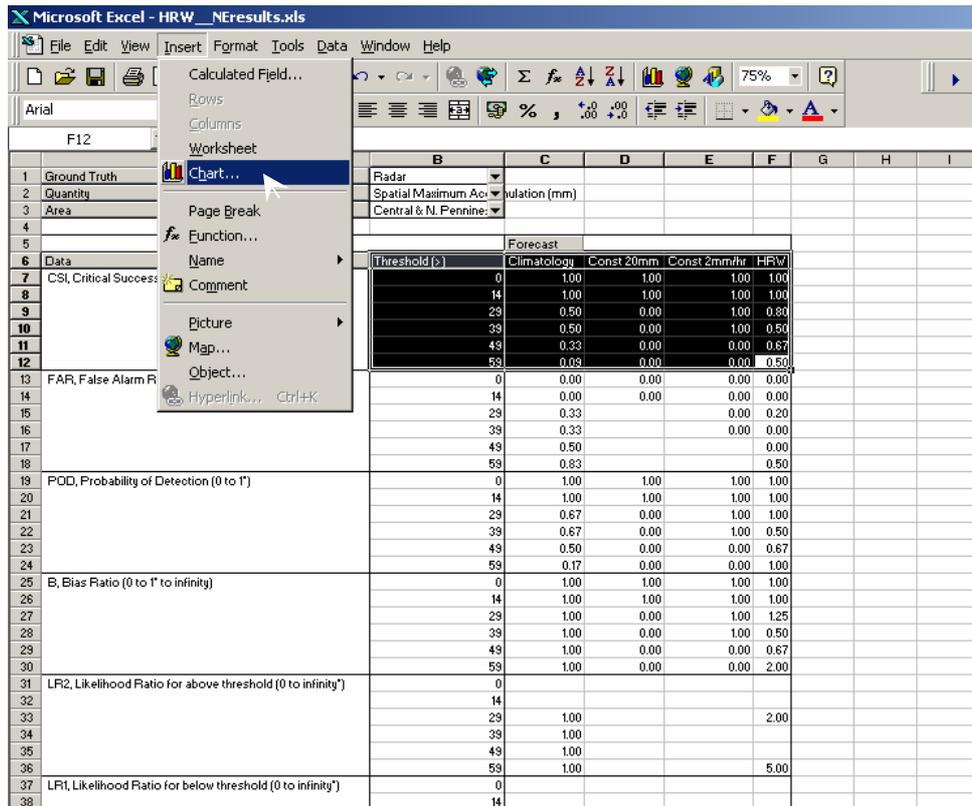
5.5 Charts

The method of generating charts from pivot tables depends on the version of Excel used. Generation of charts in Excel 97 is described in Section 5.5.1 and generation of charts in Excel 2000/2002 is described in Section 5.5..

5.5.1 Charts in Excel 97

In Excel 97, the process of generating a chart from a pivot table is the same as generating a chart from any other table of data. Figure 5.5.1 illustrates an example of a chart being generated for the CSI Skill Score. To simplify the appearance of the charts generated, it is advisable to select just a single performance measure or summary statistic to be displayed in the chart. The most recommended chart type to use is the Column chart type, subtype 1.

Step 1. With the threshold and forecast columns for a single Skill Score selected, choose "Chart" from the "Insert" menu.



Step 2. Select the Column chart sub-type 1 and click Next.

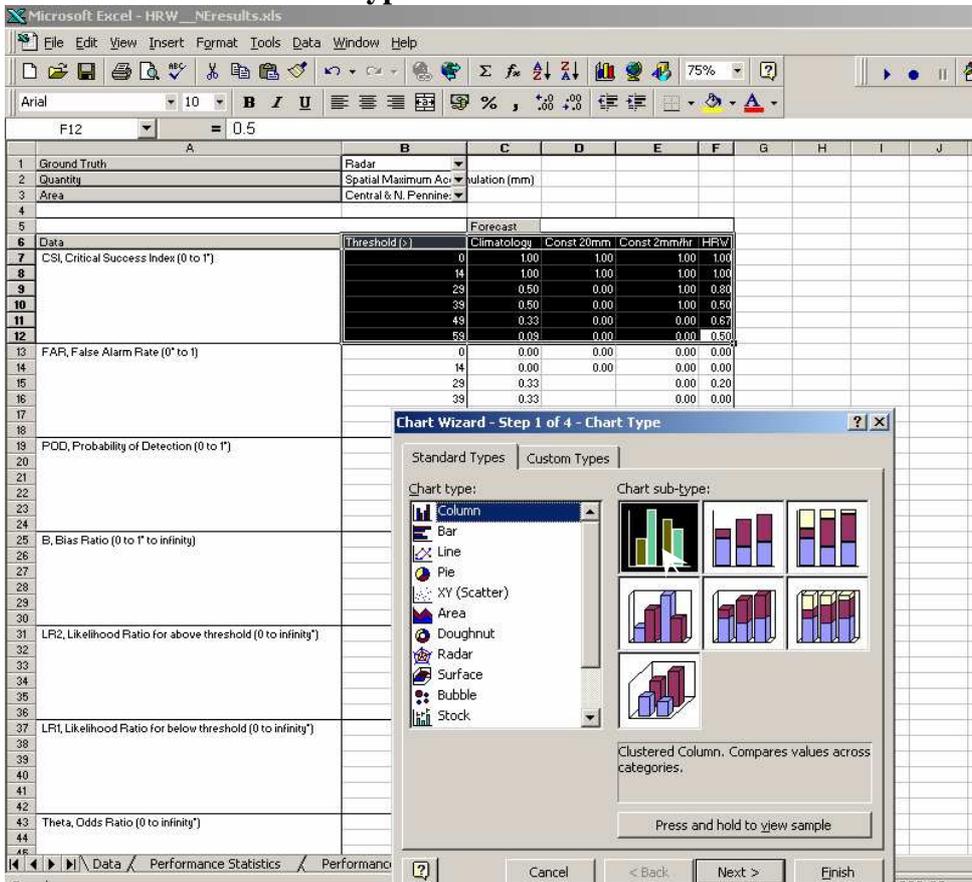


Figure 5.5.1 Generation of a chart from a pivot table in Excel 97

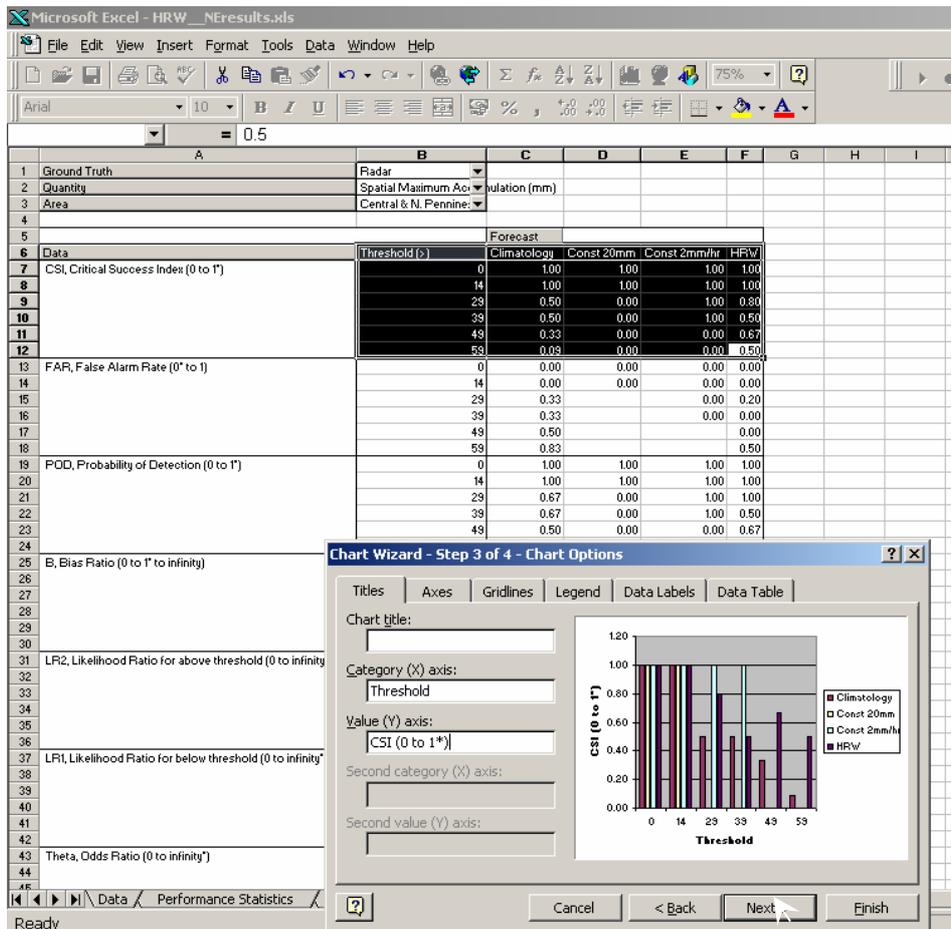
Step 3. Remove "Threshold" as a series since this should instead form the x-values.

	Threshold (>)	Climatology	Const 20mm	Const 2mm/hr	HRW
7 CSI, Critical Success Index (0 to 1)	0	1.00	1.00	1.00	1.00
8	14	1.00	1.00	1.00	1.00
9	29	0.50	0.00	1.00	0.80
10	39	0.50	0.00	1.00	0.50
11	49	0.33	0.00	0.00	0.67
12	59	0.09	0.00	0.00	0.50

Step 4. Specify the x-values for the first true series by selecting the range of threshold values on the worksheet.

Figure 5.5.1 cont' Generation of a chart from a pivot table in Excel 97

Step 5. Add an appropriate x-axis label and title.



Step 6. Embed the chart in the Skill Score Summary worksheet.

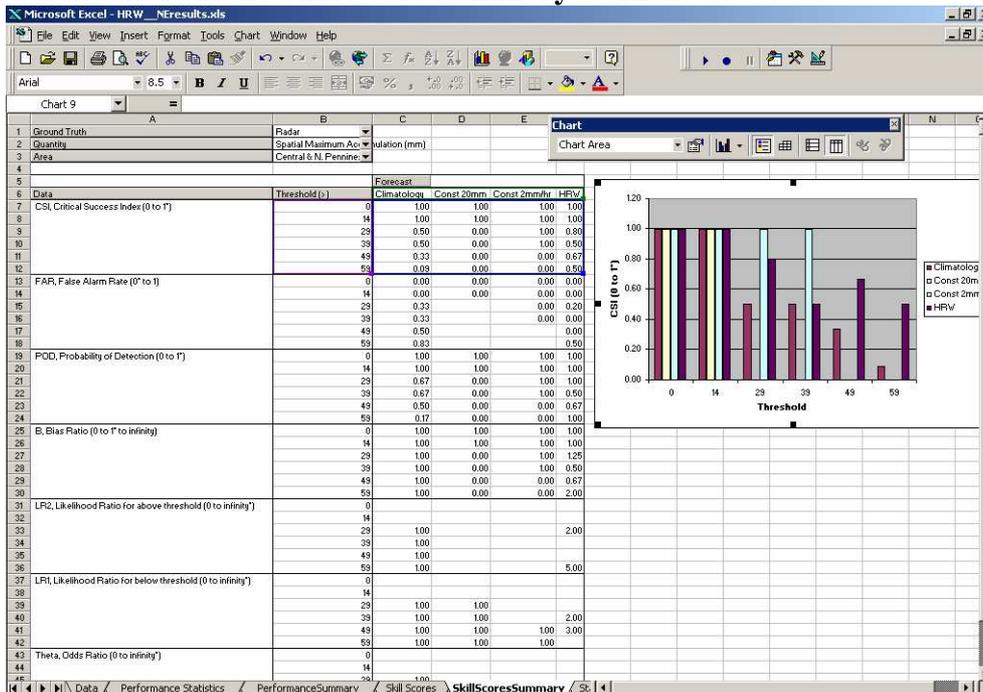


Figure 5.5.1 cont' Generation of a chart from a pivot table in Excel 97

5.5.2 Charts in Excel 2000/2002

Figures 5.5.2 (a) and (b) illustrate the process of generating a chart from a pivot table in Excel 2000/2002. The chart is generated automatically by clicking the chart button on the pivot table toolbar. The most suitable chart type to use is the Column chart type, subtype 1. You may wish to remove some items from the pivot chart to make it easier to view. Note however that the chart and table are linked so items removed from the chart are also removed from the table! These items can be put back by dragging the name from PivotTable Field List to the data area of the chart (Excel 2002) or dragging the names onto the data area from the Pivot Table Toolbar (Excel 2000). Note that when items are added to the pivot table, Excel may carry out some undesirable reformatting of the names and numbers displayed. See the Excel online help on pivot tables for more information.

Step 1. With the top left cell of the pivot table activated, click the Chart Wizard button on the PivotTable toolbar. A chart is automatically created on a new worksheet. For Steps 2 and 3 see Figure 5.5.2.

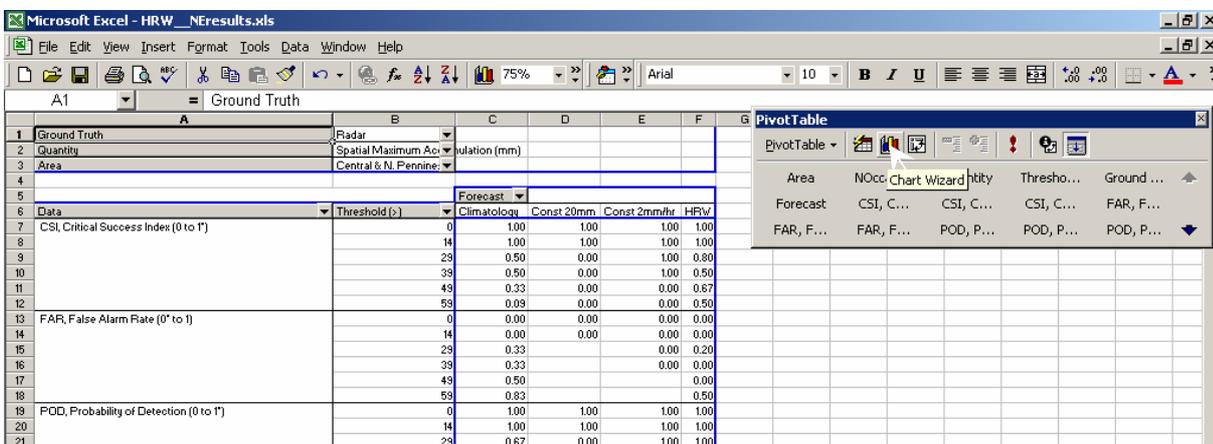
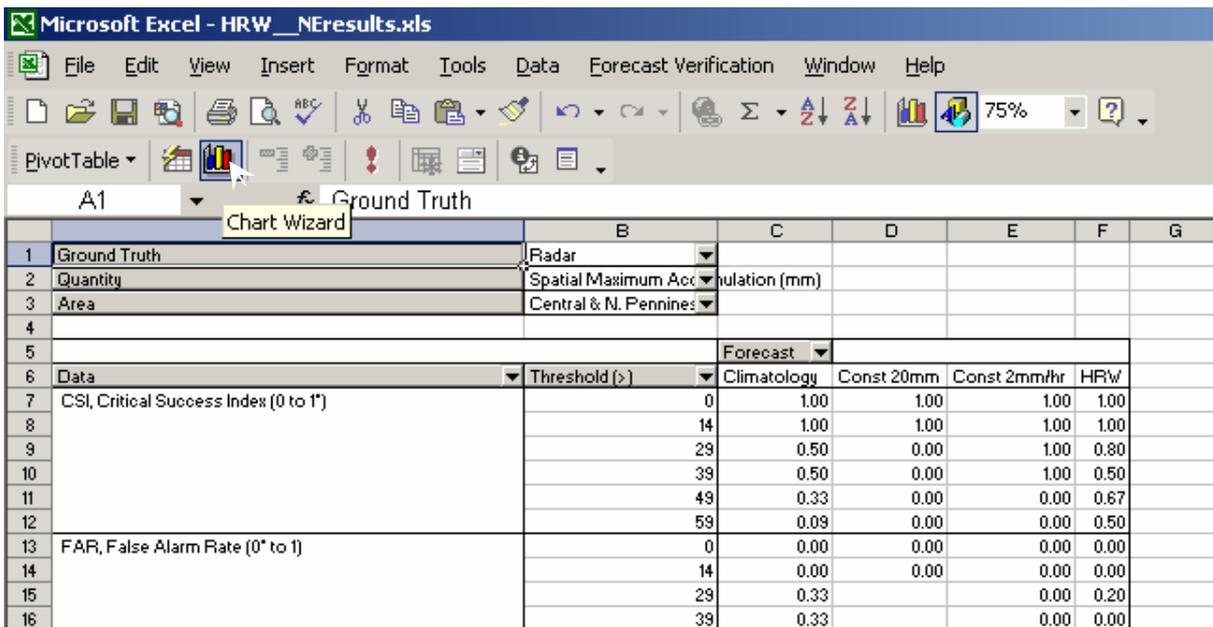


Figure 5.5.2 (a) Generation of a pivot chart in Excel 2000

Step 1. With the top left cell of the pivot table activated, click the Chart Wizard button on the PivotTable toolbar. A chart is automatically created on a new worksheet.



Step 2. Reformat the chart as required. To change the chart to a more suitable chart type, select Chart | Chart Type, and choose the Column chart sub-type 1. You may also wish to embed the chart in the pivot table worksheet (Chart | Location).

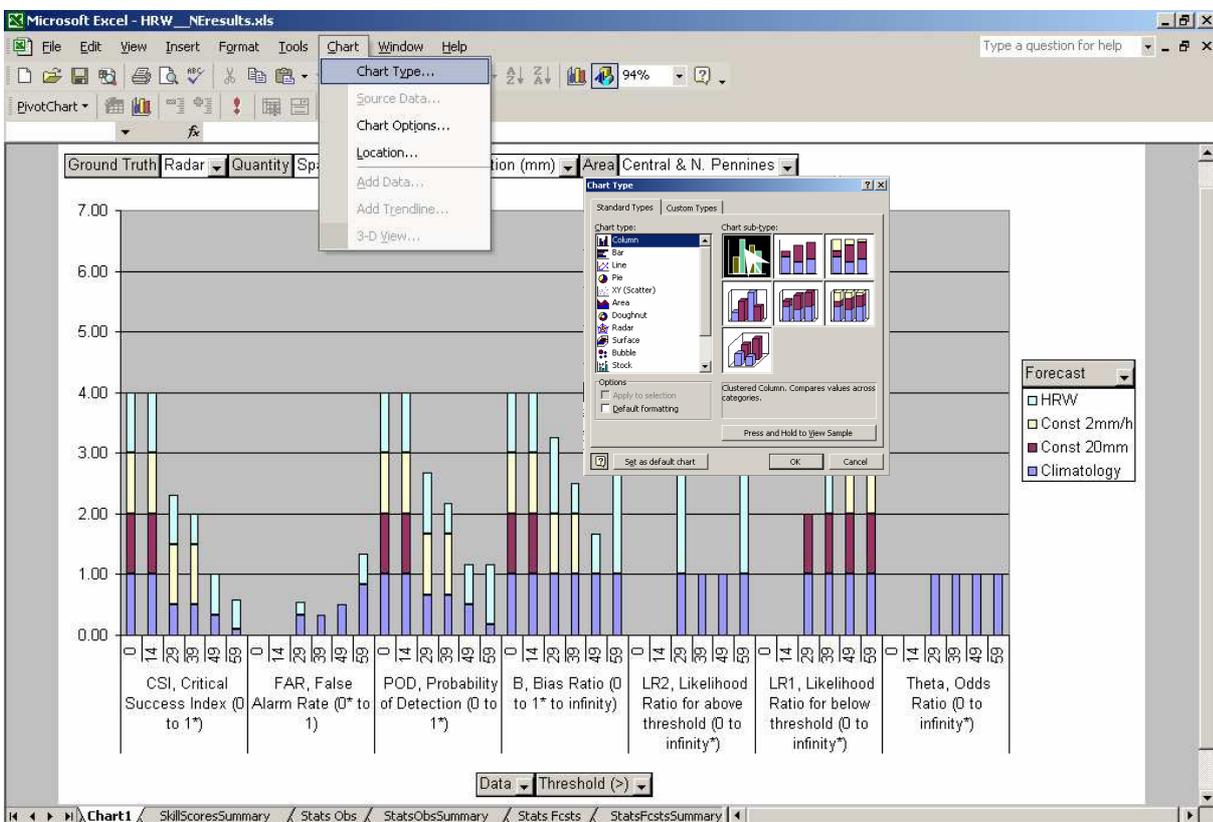


Figure 5.5.2 (b) Generation of a pivot chart in Excel 2002

Step 3. Add or remove fields from the chart using the Data drop down menu. You may need to adjust the Y-axis once you have done this.

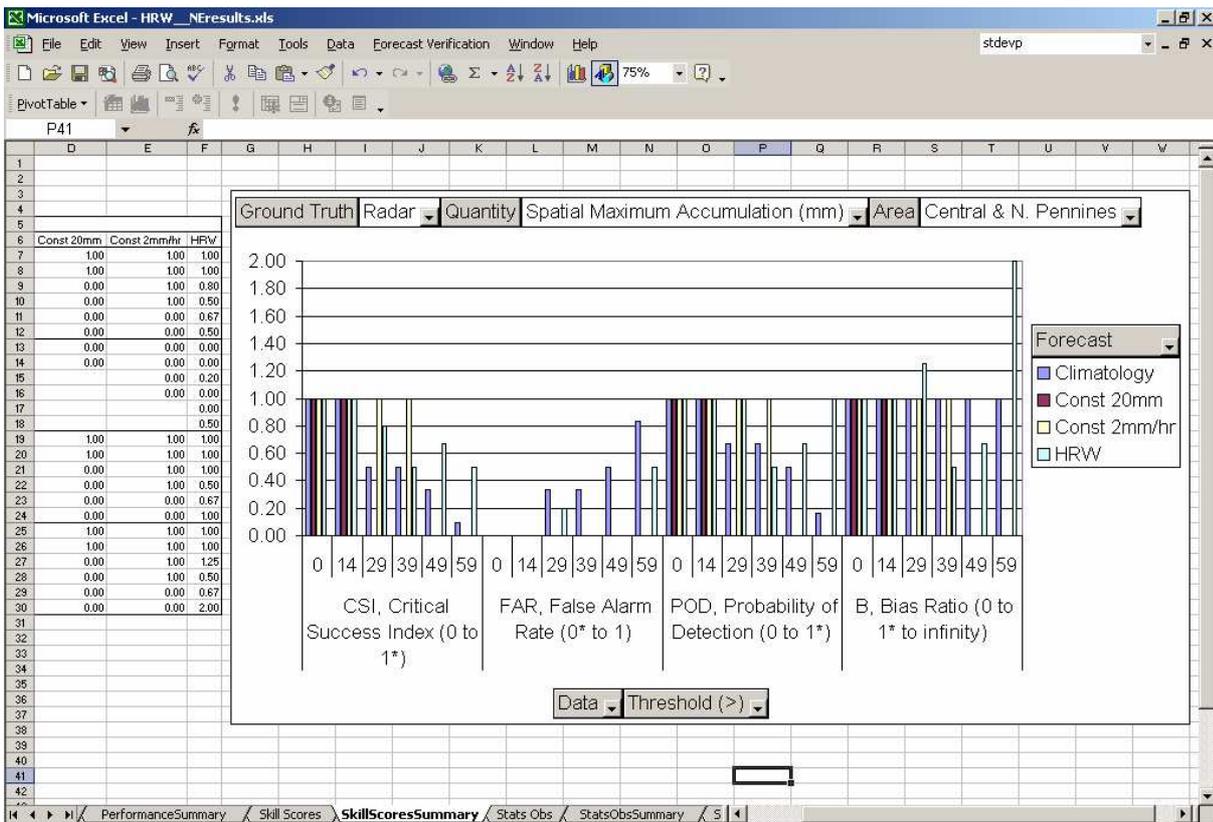


Figure 5.5.2 (b) cont' Generation of a pivot chart in Excel 2002

6 Guide to Performance Measures

An overview of the performance measures and statistics generated by the HRW Assessment Tool is given in Table 6 (a). The table gives a description of each statistic and its location in the Results Workbook. The individual performance measures along with a brief description of each are given in Table 6 (b).

Section 6.1 presents a basic guide to assessment, covering the choice of ground-truth, comparative forecasts and routine for assessment. Section 6.2 contains a guide to a selection of the most commonly used performance measures with examples for each. Section 6.3 contains an overview of the performance measures not covered in Section 6.2. A guide to comparing different sources of forecasts and ground-truths is presented in Section 6.4.

Table 6(a) Overview of statistics generated by the HRW Assessment Tool

Type of Statistic	Configuration Name	Details	Results Table Sheet Name	Summary Sheet Name
Basic statistics of observation and forecast values	Statistics of observations	Mean, median and standard deviation of observations.	Stats Obs	StatsObsSummary
	Statistics of forecasts	Mean, median and standard deviation of forecasts.	Stats Fcsts	StatsFcstsSummary
Relative Performance Measures	Typical size of error	Mean absolute error and root mean square error.	Performance Statistics	PerformanceSummary
	Bias	Mean, median and percentage errors.		
Non-continuous Performance Measures	Efficiency	R ² statistic.		
Categorical Performance Measures	Skill Scores	Categorical performance measures	Skill Scores	SkillScoresSummary
Performance Measures for comparison of forecasts	Compare forecasts	Standardised differences of Typical size of error measures across different forecasts.	Cmp Forecasts	CmpFcstsSummary
Performance Measures for comparison of observations	Compare ground-truths	Standardised differences of Typical size of error measures across different ground-truths.	Cmp G truths	CmpGSummary
Performance Measure for Probability Forecasts	Brier Score (Probability)	Continuous Brier Score for probability forecasts.	Brier Score	BrierSummary

Table 6(b) Overview of performance measures

Continuous variable:

Bias (mean error)	over- or under-estimation
Median error	over- or under-estimation
Mean absolute error	typical size of error
Root mean square error	typical size of error
R ² (Efficiency)	size of error relative to a naive forecast (equal to the mean)
% Error in maximum obs	size of error for the “biggest” event

Categorical variable:

Critical Success Index	balanced measure of forecast success
False Alarm Rate	emphasises events incorrectly forecasted
Probability of Detection	emphasises events correctly forecasted
Bias Ratio	too many or too few events forecasted
Likelihood Ratios	measure of information provided by having forecast service, separately for events and non-events
Odds Ratio	overall measure of information provided by having forecast service

Where probability forecasts are analysed:

Brier Score	error in probability terms
Continuous Brier Score	balanced measure of location and spread of forecasts relative to outcome

6.1 Guide to Assessment of Heavy Rainfall Warnings

The Tool’s assessment procedures for Heavy Rainfall Warnings are aimed at answering the basic question:

What is the typical size of error in rainfall forecasts, or rate of success in forecasting high rainfalls?

The Tool can also be used to monitor changes in forecast performance over time.

The Assessment Tool provides facilities for comparing forecasts from a number of sources using a number of different versions of ground-truth. In the following sections, we discuss the selection of target quantities and of suitable forecasts and ground-truths.

6.1.1 Selecting Target Quantities for Assessment

Use of the HRW Assessment Tool requires that the target quantities for the forecasts can be properly identified, so that suitable ground-truths can be selected, evaluated by the user and entered into the Tool. Target quantities should normally be chosen to make best use of the information given in warnings, and are therefore usually tailored to the local type of warning. If more than one type of information is provided on the warning, then there should normally be one target quantity to match each piece of information provided within a warning.

The target quantity of a forecast may be an average for an area, or a maximum within an area. Similarly, the target may be a total rainfall for a time-period (or equivalently an average rate over a time-period), or a maximum rate within a time period. Thus, typical examples of target quantities would be maximum rainfall accumulation, mean rainfall accumulation and maximum rainfall rate.

Where the target quantities relate to “maxima”, difficulties arise over defining this concept in a way that gives a meaningful value, taking into account the effect of using data of differing spatial or temporal resolution. There can also be questions over defining a “typical” rainfall value if the average value (specifically the mean value) is not quite what is required.

Examples of selection of target quantities based on representative types of warnings are given in Section 7.

6.1.2 Selecting and Determining Ground-truths for Assessment

For each target quantity, suitable ground-truth values need to be collected. Typically these will be radar data and/or raingauge network data.

If ground-truth is obtained from a raingauge network, then these data should be adequately quality-controlled. Procedures for simple inter-gauge comparisons are required. Values of ground-truth should be prepared according to whatever interpretation of “ground-truth” is acceptable, bearing in mind any discussions between the Met Office and the Environment Agency concerning what targets of the forecast should be. For merged radar-raingauge products the possible effects of incorrect raingauge data having been used would need to be considered.

In principle, merging of information from raingauge networks and weather radar should be the best source of ground-truth but this is problematic at present. Spatial averages may be adequately estimated by raingauges alone, by (adjusted) radar alone, or by use of a fully merged product. Theoretically, spatial maxima would be best estimated using radar data because of the higher spatial resolution, but experience has shown that forecasts of maxima may be better matched to the maxima obtained from a raingauge network.

The Nimrod “merged” product does not yet exist to provide experience on which advice can be based and the Nimrod “Quality Controlled” product is still undergoing changes and development. The suitability of locally-archived Nimrod data for post-event analyses has not yet been assessed: there may be a need for post-event quality control of Nimrod data. Similar problems arise for other radar-raingauge products constructed for real-time use.

Note that the time period used to calculate the ground-truths should be selected to match the time period given in the warning. The examples shown in Section 7 below contain examples of how the period of the event should be estimated from various forms of warning. Section 7 also gives examples of the practical selection and use of ground-truths. Similarly the level of spatial averaging should be tailored as closely as possible to the target quantities.

6.1.3 Selecting Forecasts

The primary forecast(s) will usually be obtained from the operational Heavy Rainfall Warnings. However, in order to assess how good an operation warning is, the forecast can be compared against selected naive forecasts. Naive forecasts are generally constructed according to simple rules, and they provide a performance baseline. One naive forecast is to only ever forecast a fixed amount, say 20mm. In this case, we would construct a forecast that was for 20mm whenever a warning is issued. Use of this naive forecast allows comparison of whether the operation forecast values are better or worse than the much simpler naive forecast. Another common naive forecast can be constructed by forecasting an amount which is proportional to the length of the event (as contained in the Heavy

Rainfall Warning). For example, one could forecast at a rate of 2mm/hr for the duration of the event, so for a 9 hour event, we forecast 18mm.

The aim of naive forecasts is really to allow assessment of the performance of operational warnings. If operational warnings do no better than naive forecasts, then we can conclude that the operational warnings are not doing particularly well and it might be appropriate to make allowance for this when interpreting warnings.

Selection of the actual values at which to set the naive values will depend on what issues are of interest. If rainfall warnings are triggered at a certain rainfall value, then this might make an appropriate choice of naive forecast. If the naive forecasts are set very low, or very high, then the HRW Tool results are unlikely to be very useful as they are likely to simply show very poor performance. The HRW tool allows for a range of constant naive forecasts to be tested, so given sufficient data, one could compare naive forecasts and determine which gives the best performance.

Examples of selection and calculation of naive forecasts in relation to specific warnings are given in Section 7.

A further type of forecast is used by the HRW Tool when assessing Skill Scores: this is the climatology forecast. The climatology forecast aims to mimic the characteristics of the observations in terms of the frequency of different sized events, but with random timing. Thus it contains the same number of exceedances and non-exceedances of the selected threshold as the observed data, but the timing of these exceedances/ non-exceedances is random. The climatology forecast is included because any reasonable forecast should outperform the climatology forecast.

6.1.4 Practicalities and limitations of using the Assessment Tool

The interpretation of individual Heavy Rainfall Warnings for use within the Tool can be problematic where the Warnings do not have a fixed structure. The Tool requires matched sets of forecast amounts or rates and corresponding ground-truths, which, in principle, requires that the Warnings be interpreted as providing quantitative forecasts for specific areas and specific time-periods. If there are missing values amongst the ground-truths and forecast values for a warning at particular site, then this warning will be excluded from the analysis. It is necessary to exclude warnings with missing values because cross-comparison between different forecasts and ground-truths could otherwise easily give very misleading results. Note particularly that because missing values will lead to data from a warning being excluded, it is important to balance the desire to compare different ground-truths and forecasts, against the desire to make the most use of the available data. In general it is not sensible to include extra ground-truths or forecasts if there are several missing values. If it is necessary to use a ground-truth or forecast that contains several missing values, it may be best to exclude them from the main analysis and then to carry out a separate analysis in which they are included.

The HRW Assessment Tool is designed for the situation where entry of all data required will be accomplished manually. This has affected the choice of assessment procedures, leading to the adoption of a forecast-by-forecast based assessment procedure. This means that the Tool's assessment procedure measures how well the forecasts of rainfall contained in the Warnings perform in matching the eventual outcomes. The HRW Tool and methods are only designed to handle analysis of warnings that have been issued. In particular, the procedures used by the tool do not allow for assessment of non-issued (missed) warnings: it would require a different approach (e.g. continuous assessment) and would need substantially more data than can be handled

conveniently using manual data-entry. More detail on the choice of methodology and assessment procedures can be found in the technical report on which the HRW Tool is based (Jones et al., 2003).

Note also that when using Skill Scores, use of a comparison threshold that is too close to the ends of the range of values occurring or forecast will not give useful results. A consequence of this is that you cannot use the Tool to assess the performance of Warnings against the threshold level criteria used for their issue. The Skill Scores analysis is most useful for assessing threshold levels where the chance of a warning being missed at this level is small. In practice, warnings should not often be missed because even if a HRW is not forecast in advance, it will usually be issued after the start of an event. If a user chooses thresholds of reasonable size somewhere in the middle of the range of observed and forecast values, this should provide useful information about the skill of the forecasts particularly when compared with the naïve and climatology forecasts.

Many of the assessment statistics are not applicable to the single warning case, and only a limited amount of information will be produced in such cases. In particular, skill scores and forecast / ground-truth comparisons cannot be provided for the single site case. Note that where it is not possible to calculate a statistic because only a single warning is available, then the pivot tables will show a blank entry for this statistic.

The Tool is primarily designed for use where there are multiple warnings available. As described above, single warnings do not provide a useful basis for comparison of forecasts. The more warnings there are the better the basis for comparison. As a very rough practical guide, 5 or more warnings should be considered the minimum for informed comparison.

6.2 Guide to Performance Measures, Part 1

6.2.1 Example

For the purposes of illustrating the performance measures, the example below has been used.

A set of 5 forecast Heavy Rainfall Warnings of the Spatial Maximum Accumulation (mm) are to be assessed for Northeast Region South Pennines Area. Radar data (Nimrod QC 2km) have been selected to provide the ground-truth. Radar has the potential to provide a better spatial maximum rainfall estimate than use of data from a raingauge network. (Note that this may not be the case in practice due to problems with radar rainfall estimation.) The values concerned are tabulated below.

Table 6.1 Example data for illustrating performance measure and skill scores

Start Time/Date	Forecast HRW	Ground-truth Radar
15:00 29-Jul-02	30	189.88
15:00 30-Jul-02	60	102.78
03:00 1-Aug-02	60	46.47
08:00 4-Aug-02	15	34.09
06:00 9-Aug-02	30	51.88

6.2.2 Guide to Notation

y_i is the observed (ground-truth) value of rainfall for sample i ($i=1,2,\dots, n$).

\hat{y}_i is the forecast value of rainfall for sample i .

Summation operator, \sum

$$\sum y_i \equiv \sum_{i=1}^n y_i = y_1 + y_2 + \dots + y_n$$

Example: Mean of ($n=5$) radar observations

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum y_i = \frac{1}{5} \sum_{i=1}^5 y_i = \frac{1}{5} (y_1 + y_2 + y_3 + y_4 + y_5) \\ &= 0.2 (189.88+102.78+46.47+34.09+51.88) = 85.02 \text{ mm} \end{aligned}$$

6.2.3 Continuous Performance Measures

The performance measures that are describe in this section are presented on the performance measures summary sheet. An example of performance measures output for the data shown in Table 6.1 is presented in Figure 6.2.

Ground Truth	Radar			
Quantity	Spatial Maximum Accumulation (mm)			
		Forecast Name		
Data	Area	const 50mm	HRW	
Mean Absolute Error	S. Pennines	42.80	51.43	
Root Mean Square Error	S. Pennines	67.26	75.39	
Mean Error (Bias)	S. Pennines	35.02	46.02	
% Error in Maximum Obs (%)	S. Pennines	73.67	84.20	
Median Error	S. Pennines	1.88	21.88	
R-squared (dimensionless)	S. Pennines	-0.37	-0.72	

Figure 6.2 The performance measure summary sheet for the data presented in Table 6.1. The HRW forecast column shows the performance statistics based on the Heavy Rainfall Warnings data. A constant naive forecast of 50mm is also shown.

Bias (mean error)

Mean of the rainfall forecast errors.

$$bias = n^{-1} \sum (y_i - \hat{y}_i)$$

Use: Indicates over-estimation (negative) or under-estimation (positive) of rainfall forecast.

Example: Bias of Heavy Rainfall Warning forecasts

Forecast errors, $y_i - \hat{y}_i$, are: 159.88, 42.78, -13.53, 19.09, 21.88.

$$\text{bias} = n^{-1} \sum (y_i - \hat{y}_i) = 0.2 (159.88 + 42.78 - 13.53 + 19.09 + 21.88) = 46.02 \text{ mm}$$

This indicates that forecasts typically underestimate by 46 mm.

Median error bias

Median of the rainfall forecast errors.

50% point of errors

Use: Indicates over-estimation (negative) or under-estimation (positive) of rainfall forecast.

Example: Median error of Heavy Rainfall Warning forecasts

Forecast errors, $y_i - \hat{y}_i$, ranked in order of size are: 159.88, 42.78, 21.88, 19.09, -13.53

The Median Error is given by the 50% point of errors, which is 21.88 mm.

This indicates forecast underestimation by 22 mm, whilst mean error bias is 46 mm.

The median error as a measure of bias is *more robust to outliers* than the mean error, giving a more typical bias value in this case.

Mean absolute error

Mean of the absolute values of the rainfall forecast errors.

$$\text{mae} = n^{-1} \sum |y_i - \hat{y}_i|$$

Use: Typical size of rainfall forecast error.

Example: Mean absolute error of Heavy Rainfall Warning forecasts

Absolute value of forecast errors, $|y_i - \hat{y}_i|$, are: 159.88, 42.78, 13.53, 19.09, 21.88.

$$\text{mae} = n^{-1} \sum |y_i - \hat{y}_i| = 0.2 (159.88 + 42.78 + 13.53 + 19.09 + 21.88) = 51.432 \text{ mm.}$$

Root mean square error

Square root of the mean of the squared rainfall forecast errors.

$$\text{rmse} = \sqrt{n^{-1} \sum (y_i - \hat{y}_i)^2}$$

Use: Typical size of rainfall forecast error.

Example: Root Mean Square Error of Heavy Rainfall Warning forecasts

Square of forecast errors, $(y_i - \hat{y}_i)^2$, are: 25562, 1830, 183, 364, 479.

$$\text{rmse} = \sqrt{n^{-1} \sum (y_i - \hat{y}_i)^2} = \sqrt{0.2 (25562 + 1830 + 183 + 364 + 479)} = 75.39 \text{ mm.}$$

Compare the *typical size of error* given by *mae* of 51.432 mm with the *rmse* value of 75.39 mm. The *rmse* is more sensitive to outliers, as seen in this example where the value calculated is inflated

by taking the square of the single large error value of 159.88. The *rmse* is arguably less *typical* than the estimate provided by the *mae* estimator.

Percentage error in Maximum observation

The error in the largest observation, as a percentage.

$$\%error = 100 \frac{(y_{\max} - \hat{y}_{\max})}{y_{\max}}$$

where y_{\max} is the maximum observation, and \hat{y}_{\max} is the forecast value for this observation.

Example: The maximum observation of the 5 Radar values is 189.88, and the corresponding forecast value is 30. Thus

$$\%error = 100 (189.88-30)/189.88 = 84.2\%$$

6.2.4 R^2 Efficiency

Proportion of variance in observations accounted for by forecast.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

$\bar{y} = n^{-1} \sum y_i$ is the sample mean of the observations

Use: Size of error relative to a simple (sample mean) forecast (dimensionless)

Example: R^2 Efficiency of Heavy Rainfall Warning forecasts

The sample mean of the radar observations has previously been calculated as $\bar{y} = 85.02$ mm.

The observed deviations from the mean, $y_i - \bar{y}$, are: 104.86, 17.76, -38.55, -50.93, -33.14.

The sum of squares of these deviations is

$$\sum (y_i - \bar{y})^2 = 10996 + 315 + 1486 + 2594 + 1098 = 16489.$$

The term $\sum (y_i - \hat{y}_i)^2$ is obtained from the *rmse* value of 75.39 previously calculated, by squaring and multiplying by 5 to give 28418. Then:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} = 1 - (28418/16489) = -0.72.$$

An R^2 of 1 indicates that we have a perfect forecast, and an R^2 of 0 means that we would do just as well if we had used a naive constant forecast that always predicted the mean value (note that this naive forecast is not practical because the mean of the radar observations is not known in advance). A negative R^2 means that the forecast is doing worse than the mean, so is a poor forecast.

6.2.5 Categorical Skill Scores

Categorical skill scores can be used to provide information about how well forecasts perform for different sizes of rainfall event. For example, we might find that events bigger than 20mm are well forecast, but events larger than 50mm are not well forecast. The Skill Scores are defined from information contained in a Contingency Table constructed from a set of forecasts and observations provided for assessment. This Contingency Table is defined below.

Contingency table:

Event Forecast	Event Observed		Total
	Yes	No	
Yes	a hit	b false alarm	$a+b$
No	c miss	d correct rejection	$c+d$
Total	$a+c$	$b+d$	$n=a+b+c+d$

An Event is defined as an exceedence of a rainfall threshold value.

a , b , c and d are the number of entries in each Event category for n rainfall forecasts and their corresponding observations.

If forecasts are provided in terms of probability tables, then the HRW Tool requires that the skill thresholds should be the same as the probability table levels. If forecasts are for a particular quantity then the thresholds are normally selected to cover a range of event sizes (ranging from small to large; see Section 7.1 below).

Example: The performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed for the data listed in Table 6.1.

The Contingency Table for this rainfall event threshold and tabulated set of rainfall values is readily calculated as:

Event Forecast	Event Observed		Total
	Yes	No	
Yes	1 hit	1 false alarm	2
No	2 miss	1 correct rejection	3
Total	3	2	5

This indicates that there are 3 observed events exceeding the 49 mm threshold, of which 1 is correctly forecast (a hit) and 2 are missed, whilst there is 1 false alarm and 1 correct rejection of an event.

The Skill Scores calculated for the above example are shown in Figure 6.3, illustrating how the results are presented on the skill scores summary worksheet. Each skill score is described below including its calculation and interpretation.

Ground Truth	Radar				
Quantity	Spatial Maximum Accumulation (mm)				
Area	S. Pennines				
			Forecast		
Data	Threshold (>)		Climatology	const 50mm	HRW
CSI, Critical Success Index (0 to 1*)	49		0.43	0.60	0.25
FAR, False Alarm Rate (0* to 1)	49		0.40	0.40	0.50
POD, Probability of Detection (0 to 1*)	49		0.60	1.00	0.33
B, Bias Ratio (0 to 1* to infinity)	49		1.00	1.67	0.67
LR2, Likelihood Ratio for above threshold (0 to infinity*)	49		1.00	1.00	0.67
LR1, Likelihood Ratio for below threshold (0 to infinity*)	49		1.00		0.75
Theta, Odds Ratio (0 to infinity*)	49		1.00		0.50

Figure 6.3. Skill scores summary data for the data presented in Table 6.2. The HRW column shows the result based on the predictions contained in the warnings. The climatology column shows the results when a randomized forecast is made for which there are the same number of values above and below the threshold as for the observed data (Section 6.1.3). The “const 50mm” column shows the results for a naive forecast of 50mm. Blank entries indicate that it is not possible to calculate a particular statistic for the given forecast. For this particular dataset the HRW performs quite poorly – overall it is outperformed by the other 2 forecast types.

Critical Success Index (Threat Score), CSI

Number correct (hits) divided by number forecast and/or observed (the threat: a+b+c)

$$CSI = \frac{a}{a + b + c}$$

Use: Balanced measure of forecast success.

Example: CSI performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$CSI = \frac{a}{a + b + c} = \frac{1}{1 + 1 + 2} = 0.25.$$

A CSI of 1 would mean the forecast is doing very well, while a score of 0 indicates that the forecast is not detecting anything correctly. A value of 0.25 indicates that the forecast is under-detecting events.

False Alarm Rate, FAR

Proportion of forecast events that fail to materialise.

$$FAR = \frac{b}{a + b}$$

Use: Emphasises events incorrectly forecasted.

Example: FAR performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$FAR = \frac{b}{a + b} = \frac{1}{1 + 1} = 0.5.$$

This can be interpreted as saying that about half of the forecast events fail to materialize.

Probability of Detection (Hit Rate for observed ‘yes’), POD

Proportion of occasions when an event does occur that are forecasted to experience the event.

$$POD = \frac{a}{a+c}$$

Use: Emphasises events correctly forecasted.

Example: *POD* performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$POD = \frac{a}{a+c} = \frac{1}{1+2} = 0.33.$$

In other words, we correctly forecast 1 in 3 events.

Bias Ratio, B

Ratio of “yes” forecasts with “yes” observations.

$$B = \frac{a+b}{a+c}$$

Use: Indicates too many (greater than 1) or too few events (less than 1) forecasted.

Example: *Bias Ratio* performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$B = \frac{a+b}{a+c} = \frac{1+1}{1+2} = 0.67.$$

This shows that we are tending to under-forecast.

In summary, the Skill Scores are: CSI=0.25, FAR=0.5, POD=0.33 and B=0.67. Thus there is a tendency to under-forecast, with a Bias Ratio B less than 1 and a low False Alarm Rate.

6.3 Guide to Performance Measures, Part 2

The results of the example calculations described below are also presented in Figure 6.3

6.3.1 Relative Categorical Skill Scores

Likelihood Ratios, LR_1 and LR_2

LR_2 is the Likelihood Ratio for correct forecast of an event.

$$LR_2 = \frac{a(b+d)}{b(a+c)}$$

The chance of forecasting that an event will occur when that event does happen is LR_2 of the chance of forecasting the event will occur when it actually does not.

LR_1 is the Likelihood Ratio for correct forecast of a non-event.

$$LR_1 = \frac{d(a+c)}{c(b+d)}$$

The chance of forecasting that an event will not occur when that event does not happen is LR_1 of the chance of forecasting the event will not occur when it actually does happen.

A good forecast service will have Likelihood Ratios greater than 1.

Use: Measure of information provided by having forecast service, separately for events and non-events

Example: Likelihood Ratio performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$LR_2 = \frac{a(b+d)}{b(a+c)} = \frac{1(1+1)}{1(1+2)} = 0.67.$$

The chance of forecasting that the event will occur when the event does happen is $2/3$ of the chance of forecasting the event will occur when it actually does not.

$$LR_1 = \frac{d(a+c)}{c(b+d)} = \frac{1(1+2)}{2(1+1)} = 0.75.$$

The chance of forecasting that the event will not occur when the event does not happen is $3/4$ of the chance of forecasting the event will not occur when it actually does happen.

We can conclude that for this example, the forecasting is not doing particularly well because the likelihood ratio is less than 1.

Odds Ratio, θ

Compares the conditional odds of making a good forecast (a hit) to those of a bad forecast (a false detection).

The *odds* (or risk) Ω of an event is the ratio of the probability p of it occurring to it not occurring, $1-p$, and so $\Omega = p/1-p$.

$$\theta = \frac{ad}{bc} = LR_1 LR_2.$$

Odds of an observed event being correctly forecast is the *Odds Ratio* times the odds of a no-event being forecast as an event.

Use: Overall measure of information provided by having forecast service.

Example: Odds Ratio performance of the Heavy Rainfall Warning forecast in warning of rainfall events in excess of 49 mm is to be assessed. Using the Contingency Table entries gives:

$$\theta = \frac{ad}{bc} = \frac{1 \times 1}{1 \times 2} = 0.5.$$

Alternatively, the product of the two Likelihood Ratios, 0.67 times 0.75, gives the same result.

Thus, the odds of an observed event being correctly forecast is *half* the odds of a no-event being forecast as an event. A good forecast service has an *Odds Ratio* greater than 1.

For this example, an odds ratio of 0.5 suggests that the forecasting is not doing particularly well because the odds ratio is less than 1.

6.3.2 Skill Scores for Probability Forecasts

Brier Score (Categorical)

Mean square probability error.

$$BS = n^{-1} \sum (Y_i - \hat{Y}_i)^2$$

Y_i indicator of event $y_i \leq x$ in the observed sample,
equal to 1 if event $y_i \leq x$ does occur, 0 if not

\hat{Y}_i probability of event $y_i \leq x$ occurring,
as stated in the probability forecast, value in the range 0 to 1

Here y_i is the observed value of sample i ($i=1,2,\dots, n$), and x is a threshold value defining the categories of event-occurrence or non-occurrence

Use: Typical size of error in probability terms.

Brier Score (Continuous)

Integrated mean square probability error.

$$BS = n^{-1} \sum \int (Y_i(x) - \hat{Y}_i(x))^2 dx$$

$Y_i(x)$ indicator of event $y_i \leq x$ in the observed sample,
equal to 1 if event $y_i \leq x$ does occur, 0 if not

$\hat{Y}_i(x)$ probability of event $y_i \leq x$ occurring,
as stated in the probability forecast, value in the range 0 to 1

Here y_i is the observed value of sample i ($i=1,2,\dots, n$), and x is a variable threshold value covering all possible values of rainfall amount or rate.

Use: Balanced measure of location and spread of rainfall forecasts relative to outcome.

6.4 Guide to Making Comparisons with the HRW Assessment Tool

6.4.1 Guide to Comparing Forecast Sources

The value of naive forecasts for comparison against the operational forecasts was indicated in Section 6.1. Good operational forecasts should out-perform naive forecasts. This leads to the question of comparing the performances of forecasts from different sources. When sample sizes are small, or when there is little difference in performance, any apparent difference may be due entirely to chance. The *standardised difference* of a performance measure for two forecast sources is used to indicate the extent of evidence that one source of forecasts is better than another.

The standardised difference is available for performance measures of the basic form

$$P = n^{-1} \sum g(\text{outcome}, \text{forecast})$$

where n is the number of forecasts assessed and $g(\dots)$ is some error function of the forecast and outcome (ground-truth) values (for example, the forecast error squared for *rmse*-type performance measures).

Then the difference in the performance measures for two sources is the average value of the differences

$$x_i = g(\text{outcome}, \text{forecast}^{(1)}) - g(\text{outcome}, \text{forecast}^{(2)}).$$

One forecast is better than another if the long-run average of the x_i 's is different from zero. The evidence for or against one source being better than another can be quantified by the value of the standardised difference, t , where

$$t = \frac{\text{mean of } x_i \text{'s}}{\text{typical error of mean of } x_i \text{'s in estimating long - run average}}.$$

The standardised difference, t , is evaluated from the sample mean, \bar{x} , and sample variance, s^2 , of the differences, x_i , as follows:

$$\bar{x} = n^{-1} \sum x_i$$

$$s^2 = (n-1)^{-1} \sum \{x_i - \bar{x}\}^2$$

$$t = \frac{\bar{x}}{\sqrt{n^{-1}s^2}}.$$

The standardised difference should be compared against the following suggested limits to assess whether there is reasonably strong evidence that one forecast source is better than another:

± 2	if sample size n is large
± 2.1	if sample size $n = 20$
± 2.5	if sample size $n = 10$
± 3.5	if sample size $n = 5$.

6.4.2 Guide to Comparing Ground-truths

There are often several different ways in which ground-truth for rainfall quantities can be determined, particularly where the targets of rainfall forecasts is unclear. The HRW Assessment Tool can be used to make a comparison of ground-truths with the aim of assessing whether the forecasts are better matched to one version of ground-truth than another. The method for doing this is rather similar to comparing different sources of forecasts (Section 6.4.1), and a full account is not given here. Once again a standardised difference approach can be used for some types of performance measure.

The standardised difference is available for performance measures of the basic form

$$P = n^{-1} \sum g(\text{outcome}, \text{forecast}).$$

Then the difference in the performance measures for two ground-truths is the average value of the differences

$$x_i = g(\text{outcome}^{(1)}, \text{forecast}) - g(\text{outcome}^{(2)}, \text{forecast}).$$

Then the procedure for treating these x_i 's is exactly the same as in Section 6.4.1.

7 Example Assessments

This section gives 3 examples of how to set up the data required for warning assessments. The warnings and ground-truth data are provided below. The data and comments provided below correspond to the example workbooks provided on the CD as part of the HRW-Tool.

The examples are deliberately selected from specific Environment Agency regions. This allows illustration of the practical issues in interpretation of the data presented in typical warnings. The regions used here are intended to provide a representative selection of the different types of warning that are available. Users should be able to apply the principles to their own region.

7.1 Northeast Region

The following example is treated in 2 parts, with 3 initial warnings (part 1), and then a further warning in Part 2.

Users may wish to refer to the following example spreadsheets:

HRW__NE_1_ConfigPart1.xls – shows the configuration,

HRW__NE_2_DataPart1.xls - shows the configuration plus the data for the first 3 warnings,

HRW__NE_4_Datapart2.xls – shows the configuration plus the data for the final warning.

These workbooks demonstrate how the data from the warnings and the ground-truths have been translated into the worksheets.

7.1.1 Heavy Rainfall Warnings

Part 1: 29 -31 July 2002

The Heavy Rainfall Warnings issued in Part 1 of the assessment period are given below.

In order to configure that data sheet, it is necessary to decide on appropriate target quantities and ground-truths, and to select appropriate forecasts for comparison.

First we consider what the target quantities should be. We note that the first 2 warnings give possible local totals (30mm and 60mm). This suggests that a spatial maximum accumulation would be an appropriate target quantity. Warning 3 provides 2 quantities. Here, we interpret the total (40mm) as a spatial mean accumulation, and the “exceptionally 60mm” as a spatial maximum accumulation. The spatial mean accumulation provides a second target quantity to be evaluated.

Appropriate ground-truths for (i) the spatial maximum accumulation are maximum radar and maximum raingauge values, and for (ii) the spatial mean accumulation are the mean radar and mean raingauge totals for the spatial mean rainfall. The values of these are given in 7.1.2.

The warnings provide a HRW forecast value. We choose to compare these warnings with 2 naive forecasts, one of constant value, one of constant rate. A constant prediction of 20mm is selected as the first naïve forecast. An alternative naive forecast is chosen to correspond to rainfall falling at a constant rate of 2mm/hr over the period of the event. The choice of values of 20mm and 2mm/hr is relatively arbitrary and was selected so as to give a reasonable comparison with the HRW forecasts. If the naive forecasts are set very low, or very high, then the results are unlikely to be very useful.

For this example, calculation of skill scores is desired. This requires selection of a number of thresholds. In this case we choose a range of thresholds that vary from small to large events. The thresholds are set at 14, 29, 39, 49, 59. They are set to just below typical storm sizes (eg just below 40mm) in order that a 40mm event is included in the >39 threshold band. The actual selection of values is relatively arbitrary, but covers both relatively small and relatively large events.

Note that the Northeast weather warnings apply to the following 7 areas: Central & N. Pennines, Cheviot, Moors, N.E. Coast, S. Pennines, Vales & Wolds and W. Pennines.

Warning 1

WEATHER WARNING

Tel: 0845 300 0300 www.metoffice.gov.uk

Environment Agency Leeds (Ref: M037)


Page 1 of 1

Warning Issued at 06:39, Mon 29 Jul 2002

Warning of: Thunderstorms

Text: Isolated thunderstorms are expected to develop after about 1500 today and persist overnight. Local totals of 30mm or more are possible, with Vale and Wolds, South Pennines, Central and North Pennines and Moors being most at risk.

Warning Number: 147
Warning Valid from: Mon 29 Jul 2002 1500
Warning Valid until: Tue 30 Jul 2002 15:00

Please telephone 0161 932 7161 if you require further information.

The following information is extracted from this warning. The start time is 15:00 29 July, and the end time is 15:00 30 July, i.e. 24 hours later. The HRW forecast of spatial maximum rainfall is 30mm: this applies to the 4 areas listed. The naive 20mm constant forecast is 20mm, and the naive 2mm/hr constant rate forecast is 48mm as the event lasts 24 hours.

Warning 2

WEATHER WARNING	
Tel: 0845 300 0300 www.metoffice.gov.uk	
Environment Agency Leeds (Ref: M037)	Page 1 of 1
Warning Issued at 07:15, Mon 29 Jul 2002	
Warning of: Heavy Rain	
Text: Periods of heavy rain are expected to develop on Tuesday. Some large totals of rain are expected in the following 24h, possibly up to 60mm locally. All areas at risk, except, perhaps Northeast Coast and Cheviot.	
Warning Number: 148 Warning Valid from: 30 July 1500 Warning Valid until: 31 July 1500	
Please telephone 0161 932 7161 if you require further information.	

This warning is interpreted similarly to the first warning. It is taken to apply to 5 regions (ie not Northeast coast or Cheviot).

Warning 3

WEATHER WARNING	
Tel: 0845 300 0300 www.metoffice.gov.uk	
Environment Agency Leeds (Ref: M037)	Page 1 of 1
Warning Issued at 09:38, Wed 31 Jul 2002	
Warning of: Heavy Rain	
Text: Another rainfall event is expected to begin tonight. Totals of 40mm and exceptionally 60mm, are possible in the 24 hour period beginning around 0300 tomorrow. All areas are at risk.	
Warning Number: 175 Warning Valid from: Thursday 1st August 0300 Warning Valid until: Fri 02 Aug 2002 0300	
Please telephone 0161 932 7161 if you require further information.	

The following information is extracted from Warning 3. The start time is 03:00 1 August, and the end time is 03:00 2 August, i.e. 24 hours later. The HRW forecast of spatial maximum rainfall is 60mm and of spatial mean rainfall is 40mm: this applies to all 6 areas.

The naive 20mm constant forecast is 20mm and the naive 2mm/hr constant rate forecast is 48mm as the event last 24 hours. The naive forecasts are the same for both target quantities.

Part 2: 1 August 2002

The Heavy Rainfall Warning issued in part 2 of the assessment period is given below.

WEATHER WARNING	
Tel: 0845 300 0300 www.metoffice.gov.uk	
Environment Agency Leeds (Ref: M037)	Page 1 of 1
Warning Issued at 23:03, Thu 01 Aug 2002	
Warning of: Heavy Rain	
Text: Further heavy rain is expected right through Friday with 15-30mm possible in 12h periods. Initially Moors, Vale and Wolds, C+N Pennines and West Pennines will be affected, but from around midday the areas with heavy rain will transfer northwards. NE Coast and Cheviot, with the north of C+N and W Pennines will be affected then, whilst more southern areas become dry by evening.	
Warning Number: 16	
Warning Valid from: 02 Aug 2002 0600	
Warning Valid until: 03 Aug 0300	
Please telephone 0161 932 7161 if you require further information.	

For this warning, a range of values is provided. Here, we assume that the upper value (30mm) is equivalent to a spatial maximum rainfall, and decide not to use the lower (15mm) value. This interpretation is subjective. The warning states that the duration of the event is for 12 hours. The 4 regions Moors, Vales and Wold, Central and North Pennines and West Pennines are initially affected. So for these catchments, the start time is set to the start of the warning (06:00 2 August) and the end to 18:00 2 August. For NE Coast and Cheviot, the start time is taken to be midday (12:00) and the end time is set to 12 hours later.

7.1.2 Ground-truth

The radar and raingauge ground-truth values for the period 29 July - 1 August are given in Table 7.1.2.

Table 7.1.2 Ground-truth data for Northeast demonstration event. Only data used in the example workbook are included.

Period Start	Period End	Area	Maximum Raingauge (mm)	Maximum Radar (mm)	Mean Raingauge (mm)	Mean Radar (mm)
15:00 29 July	15:00 30 July	Central & N. Pennines	20.60	58.03		
		Cheviot		20.32		
		Moors	2.20	34.69		
		N.E. Coast		50.13		
		S. Pennines	35.80	189.88		
		Vales & Wolds	16.20	196.66		
		W. Pennines				
15:00 30 July	15:00 31 July	Central & N. Pennines	60.40	85.59		
		Cheviot				
		Moors	3.20	63.84		
		N.E. Coast				
		S. Pennines	68.80	102.78		
		Vales & Wolds	33.80	83.25		
		W. Pennines	30.00	62.38		
03:00 1 Aug	03:00 2 Aug	Central & N. Pennines	37	54.38	21.9	17.16
		Cheviot	3.6		1.29	.089
		Moors	73.2	56.84	51.00	32.71
		N.E. Coast	37.2		12.53	6.32
		S. Pennines	33.2	46.47	27.63	19.31
		Vales & Wolds	51.8	56.84	30.37	17.64
		W. Pennines	50	49.28	8.73	32.71
06:00 2 Aug	18:00 2 Aug	Central & N. Pennines	38.80	24.47		
		Cheviot				
		Moors	24.40	17.00		
		N.E. Coast				
		S. Pennines				
		Vales & Wolds	12.65	17.03		
W. Pennines	29.20	26.28				
12:00 2 Aug	00:00 3 Aug	Central & N. Pennines				
		Cheviot	34.60	25.30		
		Moors				
		N.E. Coast	13.80			
		S. Pennines				
		Vales & Wolds				
W. Pennines						

Note that for some of the dates, there are rainfall data but not radar data. The HRW Tool does not process warnings unless the data are complete for a region. Thus, at any site, *only the warnings with complete ground-truth information will be used in the calculations.*

7.2 Northwest Region

7.2.1 Heavy Rainfall Warnings

The Heavy Rainfall Warnings issued between 31 January and 2 February 2002 are shown below.

The basic type of information contained for this region's warning is fairly similar to that of Northeast Region. We therefore select the same 2 target quantities, the spatial maximum accumulation, and the spatial mean accumulation, and use the same ground-truths and forecast types as in 7.1.1.

The configuration and data for these warnings can be examined in the HRW__NW_1_ConfigAndData.xls workbook.

Warning 1

```
TO THE ENVIRONMENT AGENCY
FROM THE MET OFFICE (MANCHESTER)
ISSUED AT 0648
ON Thursday 31/01/2002

=====
CONFIRMATION COPY OF RAINFALL WARNING
=====

TEXT: Event A; Heavy rain from around 1100, will give up to 30mm
in areas 3,4 5 and 6.
CONFIDENCE LEVEL: 30%

VALID UNTIL: 2200

(C) CROWN COPYRIGHT 2002. ALL RIGHTS RESERVED.
```

The heavy rainfall value of 30mm is taken to be the spatial maximum accumulation. Only the start of the rainfall event is given (11:00 31 January). However the warning is given a “valid until” time (22:00) and this is used as the End time, thus giving the event a duration of 11 hours. The 20mm naive forecast is 20mm, and the 2mm/hr naive forecast is $2 \times 11 = 22$ mm. Note that no use is made of the confidence level data.

Warning 2

TO THE ENVIRONMENT AGENCY
FROM THE MET OFFICE (MANCHESTER)
ISSUED AT 2058
ON Thursday 31/01/2002

=====
CONFIRMATION COPY OF RAINFALL WARNING
=====

TEXT: Event A: Another period of heavy rain is expected to start around 0600 on Friday. Totals of up to 40mm are likely in 6-9h, with 12h totals locally 50mm. Areas most at risk are 3,4,5,6,7.

CONFIDENCE LEVEL: 60%

VALID UNTIL: 01/1800

(C) CROWN COPYRIGHT 2002. ALL RIGHTS RESERVED.

Warning 2 contains 2 predictions which span different durations. In order to deal with this, it is easiest to treat this warning as 2 separate warnings: one for an event of 40mm over 7.5h (the mean of 6 and 9 hours) and one for an event of 50mm lasting 12 hours. The values are taken to refer to spatial maximum accumulations.

Warning 3

TO THE ENVIRONMENT AGENCY
FROM THE MET OFFICE (MANCHESTER)
ISSUED AT 2058
ON Friday 01/02/2002

=====
CONFIRMATION COPY OF RAINFALL WARNING
=====

TEXT: EVENT A. - 30mm of rain is likely in the next 12 hours in exposed southwest facing catchments in Cumbria and North Lancashire.

Catchments at risk are 3,4,5,6 and 7.

10 to 15mm of rain is possible over remaining parts of Lancashire but further south only 5 to 10 mm of rain is likely in the 12 hour period.

CONFIDENCE LEVEL: 60 per cent.

VALID UNTIL: 020900

(C) CROWN COPYRIGHT 2002. ALL RIGHTS RESERVED.

This warning refers to expected values of rainfall rather than maximum rainfalls and so is taken as a measure of the spatial mean accumulation. As we are working with mean rainfall, we use the average rainfall amount where a range is given, i.e. 12.5 (= (10 +15)/2) mm, for the remaining parts of Lancashire (Areas 8-10), and 7.5 mm for the Areas 11-15. The start time for this event is taken as 20:58 and the end time is 12 hours later.

7.2.2 Ground-truth

The radar and raingauge ground-truth values for the period 31 January - 2 February are given in Table 7.2.1.

Table 7.2.2 Data for Northwest Region 31 Jan - 2 Feb 2002. Only data used in the example workbook are included.

Period Start	Period End	Area	Maximum Raingauge (mm)	Maximum Radar (mm)	Mean Raingauge (mm)	Mean Radar (mm)
11:00 31 Jan	12:00 31 Jan	West Lakes	47.4	45.1		
		Upper Eden	45.2	40.3		
		South Lakes (1)	54.5	53.9		
		South Lakes (2)	19.2	51.6		
06:00 1 Feb	13:30 1 Feb	West Lakes	45.4	53.5		
		Upper Eden	64	53.4		
		South Lakes (1)	42.4	51		
		South Lakes (2)	24	51		
		Lune	33.6	48		
06:00 1 Feb	18:00 1 Feb	West Lakes	52.2	61.9		
		Upper Eden	67.2	61.2		
		South Lakes (1)	47	61.2		
		South Lakes (2)	24.4	55.2		
		Lune	34.8	50.9		
20:58 1 Feb	08:58 2 Feb	West Lakes			3.2625	6.11
		Upper Eden			9.05	13.3
		South Lakes (1)			1.75	4.7
		South Lakes (2)			2.4	6.9
		Lune			4.914	6.2
		Wyre			1.95	2.7
		Ribble			2.75	4.2
		Douglas			0.5	0.7
		Irwell			2.6	2.8
		Glaze, Sankey and Ditton / Alt and Crossens			0.06667	0.5
		Mersey			1.7846	2.8
		Weaver and Gowey			1.7429	2.1
		Wirral			0.2	0.2

7.3 Thames Region

7.3.1 Heavy Rainfall Warnings

The quantitative elements of the Heavy Rainfall Warnings issued in July and August 2002 for Thames Region Northeast Area are shown below.

This Area provides a typical example of where predictions are supplied as a probability table.

The warnings refer to the point maximum of the rainfall, and to the point maximum rate. Suitable target quantities are thus the maximum rainfall accumulation and the maximum rainfall rate.

For the maximum rainfall accumulation, the data supplied from the warning can be used to provide 2 forecasts. The first is the “most likely point maximum” a single value. The second forecast is the “probability of rainfall amount” supplied in probability table form. Thus the configuration file needs to be set to indicate that this is the case and that there are 7 rainfall categories. Since Skill Scores are required to be calculated, it is necessary to specify the Skill thresholds. Whenever probability table data is available, the skill thresholds must be set to match the probability categories. Thus 7 Skill thresholds are used: 0, 10, 20, 40, 60, 80 and 100.

Rain gauge maxima can be determined for this region and are used as the ground-truth for this target quantity.

For maximum rainfall accumulation, two forecasts can again be extracted from the warnings: “most likely point max rate” and “probability of this rate”. The probability table categories and skill threshold are set up in a similar way as for the rainfall accumulation above. There are 2 available ground-truths for this target quantity: gauge maximum rate and radar maximum rate.

When entering the warning data into the HRW Tool data tables, the probability tables and the most likely maximum values can be directly entered into the appropriate columns of the worksheet. The start and end times, are taken from the timing of the event and timing of maximum rate data i.e. for Warning 1 below, they are 17:00 29 July to 00:00 30 July for maximum rainfall accumulation and 17:00 29 July to 22:00 29 July for maximum rainfall rate. No use is made of the uncertainty in the timings information contained in the warnings. Thus for calculation of ground-truths the period is simply taken to be defined by the start and end times given above.

All the warnings below are processed in an identical manner.

The configuration and data for these warnings can be examined in the workbook HRW__Thames_1_ConfigAndData.xls.

Warning 1

EA (Ref: MO34)

Warning Issued on Monday, 29 July 2002 at 17:30

Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	80
10 - 20	50
20 - 40	20
40 - 60	10
60 - 80	
80 - 100	
100 +	

Most likely point maximum	15
Timing of event	17-00
Uncertainty of timings	+2

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	90
10	70
20	60
50	10

Most likely point max rate	20
Time of max rate	17-22

Valid from 1730 on Monday 29 July 2002 to 2400 on Monday 29 July 2002

Warning 2

EA (Ref: MO34)

Warning Issued on Tuesday, 30 July 2002 at 21:36

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	100
10 - 20	80
20 - 40	60
40 - 60	20
60 - 80	10
80 - 100	
100 +	

Most likely point maximum	25
---------------------------	----

Timing of event 2130 to 0600
Uncertainty of timings +/-1

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	100
10	100
20	60
50	20

Most likely point max rate 15
Time of max rate 2130 to 2400

Valid from 2130 on Tuesday 30 July 2002 to 0600 on Wednesday 31 July 2002

Warning 3

EA (Ref: MO34)
Warning Issued on Wednesday, 31 July 2002 at 06:16
LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area
From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	100
10 - 20	70
20 - 40	50
40 - 60	20
60 - 80	5
80 - 100	
100 +	

Most likely point maximum 25
Timing of event 06-12
Uncertainty of timings +3

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	90
10	70
20	50
50	20

Most likely point max rate 20
Time of max rate 06-12

Valid from 0600 on Wednesday 31 July 2002 to 1200 on Thursday 01 August 2002

Warning 4

EA (Ref: MO34)
Warning Issued on Wednesday, 31 July 2002 at 12:20

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area
From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	90
10 - 20	50
20 - 40	20
40 - 60	5
60 - 80	
80 - 100	
100 +	

Most likely point maximum 15
Timing of event 12 to 24
Uncertainty of timings +/- 1

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	90
10	60
20	20
50	5

Most likely point max rate 25
Time of max rate 14 to 20

Valid from 1200 on Wednesday 31 July 2002 to 2400 on Wednesday 31 July 2002

Warning 5

EA (Ref: MO34)

Warning Issued on Saturday, 03 August 2002 at 14:24

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area
From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	80
10 - 20	40
20 - 40	10
40 - 60	
60 - 80	
80 - 100	
100 +	

Most likely point maximum 10
Timing of event 16-21
Uncertainty of timings +/-1

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	80
10	40
20	10
50	

Most likely point max rate 12
Time of max rate 16-21

Valid from 1500 on Saturday 03 August 2002 to 0900 on Sunday 04 August 2002

Warning 6

EA (Ref: MO34)
Warning Issued on Saturday, 03 August 2002 at 23:28
LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area
From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	90
10 - 20	50
20 - 40	20
40 - 60	
60 - 80	
80 - 100	
100 +	

Most likely point maximum 20
Timing of event 21-03
Uncertainty of timings +/-3

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	90
10	30
20	10
50	

Most likely point max rate 8
Time of max rate 23-02

Valid from 2300 on Saturday 03 August 2002 to 0600 on Sunday 04 August 2002

Warning 7

EA (Ref: MO34)
Warning Issued on Sunday, 04 August 2002 at 12:44
LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area
From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm) area	Probability of this amount at any location in the (%)
> 0 - 10	100
10 - 20	80
20 - 40	60
40 - 60	20
60 - 80	10
80 - 100	
100 +	
Most likely point maximum	25
Timing of event	13-21
Uncertainty of timings	

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr) area	Probability of this rate at any location in the (%)
4	100
10	100
20	70
50	10
Most likely point max rate	32
Time of max rate	13-18

Valid from 1300 on Sunday 04 August 2002 to 2100 on Sunday 04 August 2002

Warning 8

EA (Ref: MO34)

Warning Issued on Monday, 05 August 2002 at 10:42

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm) area	Probability of this amount at any location in the (%)
> 0 - 10	100
10 - 20	80
20 - 40	50
40 - 60	20
60 - 80	10
80 - 100	
100 +	
Most likely point maximum	30
Timing of event	11-22
Uncertainty of timings	2

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr) area	Probability of this rate at any location in the (%)
4	100
10	100
20	80
50	40

Most likely point max rate 32
Time of max rate 13-19

Valid from 1100 on Monday 05 August 2002 to 2200 on Monday 05 August 2002

Warning 9

EA (Ref: MO34)

Warning Issued on Wednesday, 07 August 2002 at 11:52

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	90
10 - 20	60
20 - 40	20
40 - 60	
60 - 80	
80 - 100	
100 +	

Most likely point maximum 30
Timing of event 15-09
Uncertainty of timings +/-3

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
	(%)
4	70
10	40
20	20
50	

Most likely point max rate 15
Time of max rate 15-03

Valid from 1500 on Wednesday 07 August 2002 to 0900 on Thursday 08 August 2002

Warning 10

EA (Ref: MO34)

Warning Issued on Friday, 09 August 2002 at 05:44

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
	(%)
> 0 - 10	100
10 - 20	70
20 - 40	30
40 - 60	0
60 - 80	
80 - 100	

100 +

Most likely point maximum	25 mm
Start of event	09/0600
End of event	09/2359
Uncertainty of timings	+/- 2 hrs

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
--------------------------	--

	(%)
4	100
10	80
20	50
50	10

Most likely point max rate	24 mm / hr
Time of max rate	15-22 hrs

Warning 11

EA (Ref: MO34)

Warning Issued on Saturday, 10 August 2002 at 11:52

LWCHRWNE Heavy Rain Warning for EA Thames - Northeast Area

From the Met Office London. Telephone 020 7204 7254

Probability of Rainfall Amounts

Rainfall amounts (mm)	Probability of this amount at any location in the area
-----------------------	--

	(%)
> 0 - 10	80
10 - 20	70
20 - 40	50
40 - 60	10
60 - 80	5
80 - 100	
100 +	

Most likely point maximum	25 mm
Start of event	10/1200
End of event	10/2100
Uncertainty of timings	+/- +1 hrs

Probability of at least this Rainfall Rate

Rainfall rates (mm / hr)	Probability of this rate at any location in the area
--------------------------	--

	(%)
4	90
10	70
20	50
50	20

Most likely point max rate	25 mm / hr
Time of max rate	12 to 19 hrs

7.3.2 Ground-truth

The radar and raingauge ground-truth values for the Heavy Rainfall Warnings issued for Thames Northeast area in July and August are given in Table 7.3.2.

Table 7.3.2 Ground-truth values for Thames Northeast Area

Period	Start	Period	End	Max Raingauge Accumulation (mm)	Max Raingauge Rate (mm/hr)	Max Radar Rate (mm/hr)
16:00	29 7 2002	23:00	29 7 2002	3.60	5.60	191.75
20:30	30 7 2002	05:00	31 7 2002	29.20	49.60	76.09
05:00	31 7 2002	11:00	31 7 2002	28.20	27.20	37.28
11:00	31 7 2002	23:00	31 7 2002	26.60	77.60	79.12
15:00	3 8 2002	20:00	3 8 2002	22.00	52.80	109.56
20:00	3 8 2002	02:00	4 8 2002	19.80	30.40	66.97
12:00	4 8 2002	20:00	4 8 2002	22.60	45.60	124.78
10:00	5 8 2002	21:00	5 8 2002	11.20	26.40	76.09
14:00	7 8 2002	08:00	8 8 2002	23.40	42.40	133.94
05:00	9 8 2002	23:00	9 8 2002	41.80	56.80	170.44
11:00	10 8 2002	20:00	10 8 2002	6.00	19.20	88.28

8 HRW Assessment Tool FAQ

References

Jones, A.E., Jones, D.A., and Moore R.J., 2003. Development of Rainfall Forecast Performance Monitoring Criteria. Phase 1: Development of Methodology and Algorithms. CEH Wallingford pp291.

Appendix A Probability interpretation of Relative Categorical Skill Scores

Contingency table in probability form

Event Forecast	Event Observed		Total
	Yes	No	
Yes	$p(f, o)$	$p(f, \bar{o})$	$p(f)$
No	$p(\bar{f}, o)$	$p(\bar{f}, \bar{o})$	$p(\bar{f})$
Total	$p(o)$	$p(\bar{o})$	1

$p(f, o)$: joint probability of a hit being forecast and observed (a yes/yes event)

$p(f)$: marginal probability for an event being forecasted

$p(f | o)$: conditional probability of a yes forecast given a yes observation

Overbar: signifies a no event eg. \bar{f} indicates forecast is that an event will not occur

Likelihood Ratios, LR_1 and LR_2

LR_1 is the Likelihood Ratio for correct forecast of a non-event.

$$LR_1 = \frac{d(a+c)}{c(b+d)} = p(\bar{f} | \bar{o}) / p(\bar{f} | o)$$

LR_2 is the Likelihood Ratio for correct forecast of an event.

$$LR_2 = \frac{a(b+d)}{b(a+c)} = p(f | o) / p(f | \bar{o})$$

Use: Measure of information provided by having forecast service, separately for events and non-events

Odds Ratio, θ

Compares the conditional odds of making a good forecast (a hit) to those of a bad forecast (a false detection).

The odds (or risk) Ω of an event is the ratio of the probability p of it occurring to it not occurring, $1-p$, and so $\Omega = p / 1 - p$.

$$\theta = \frac{\Omega(f | o)}{\Omega(f | \bar{o})} = \frac{ad}{bc} = LR_1 LR_2 .$$

Use: Overall measure of information provided by having forecast service.

Appendix B Glossary of terms

Bias The mean of the errors in the rainfall forecast. It indicates over-estimation (negative values) or under-estimation of rainfall by a forecast.

Bias Ratio An indicator of whether too many (greater than 1) or too few events are forecast.

Brier Score A statistical measure which provides the typical size of error in probability terms.

Climatology Forecast A forecast generated randomly with the same number of values above a threshold as in the observations.

Contingency Table A table giving the number of events that are, or are not, observed and/or forecast.

Critical Success Index This is the number of correct forecasts divided by total number of events that are either forecast or observed. A Skill score which provides a balanced measure of forecast success.

Efficiency See under R^2 Efficiency.

Event An exceedance of a rainfall threshold value.

False Alarm Rate. This is the proportion of forecast events that fail to materialise. A Skill score that emphasises events incorrectly forecasted.

Ground-truth An observed value, often derived from raingauge and/or radar data, that is used to approximate a true rainfall quantity at the ground surface.g. maximum rainfall rate.

Likelihood Ratio A statistical measure used here to measure how useful a forecast is.

Mean absolute error The mean of the absolute values of the rainfall forecast errors. In indicates the typical size of rainfall forecast error.

Median error The median of the rainfall forecast errors. It indicates over-estimation or under-estimation of rainfall by a forecast.

Naive forecast A simple forecast to be used for comparative purposes.

Odds Ratio A statistical measure that provides an overall measure of information provided by having a forecast service.

Probability of Detection The proportion of events that are correctly forecast. A Skill Score that emphasises events correctly forecasted.

Probability Table A table giving probabilities of whether an event will exceed a series of threshold sizes.

Root mean square error A measure of the typical size of rainfall forecast error.

R^2 Efficiency The proportion of variance in observations that is accounted for by a forecast.

Skill Scores Statistics that can be used to provide information about how well forecasts perform for different sizes of rainfall event.

Target Quantity A target quantity is a quantity that is forecast in the Heavy Rainfall Warnings.

Validity Time The time up until which the warning is valid, or a period of validity.