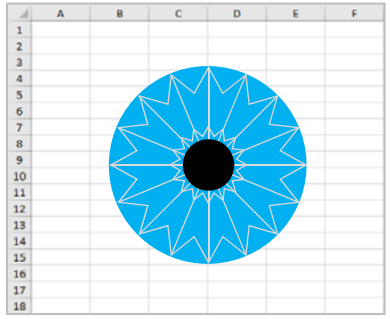
Inherent Risk In Spreadsheets (IRiS)
or
How to Keep an Eye on your Spreadsheets

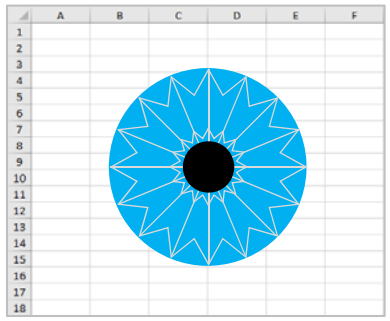
ICEAA 2016 International Training Symposium
Bristol, 17th to 20th October 2016

Alan R Jones
Estimata Limited
Promoting TRACEability in Estimating



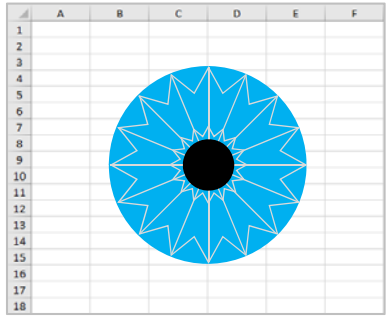
IRiS: Agenda

- Who or What is *IRiS*?
- Good Practice Spreadsheet Principles (GPS)
- A Word about Verification and Validation (V&V)
- *IRiS* Scoresheet and Chart
- Any Questions?



Who or What is IRiS?

Let's talk about IRiS



What is IRiS?

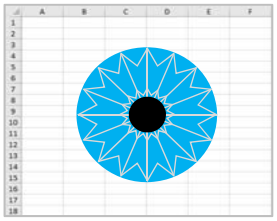
- I've known *IRiS* all her life … even before she was born
- I'm her proud father … her Dad
- … and probably her surrogate Mother as well!

- *IRiS* is not very old, less than a year, in fact, and this is her first appearance in public

- I knew what I hoped she would become when she was first conceived
- But I know that her future development will evolve based on her interaction with others. her development is not wholly in my hands

- This is your chance to contribute to the development and maturity of *IRiS*
- *IRiS* is not the real daughter I always wanted.
- *IRiS* is a visualisation tool to express the

Inherent Risk in Spreadsheets



Imagine this …

Scenario

- You've just inherited someone else's spreadsheet
- You are told to use it to inform an important business decision … tomorrow
- You open the spreadsheet and what do you find …
- The Nightmare Scenario …

  - The logic jumps all over the place
  - Some calculations are 4 lines long
  - You can't find all the input cells
  - There are no notes or comments to help
  - There's a calculation which includes a fixed exchange rate of 1.473 (but from what to what?)
  - There's cost calculations to 17 dp's
  - There are Excel functions used you never knew existed
  - There's an obvious arithmetical error that should have been found before
  - There's a macro that generates data but you don't know how it does it

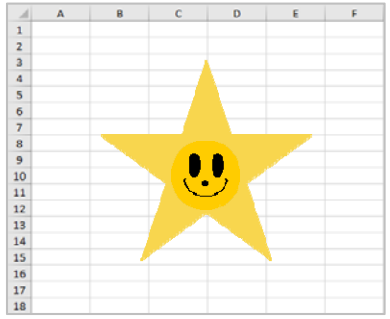
… the list could go on

Good Practice Spreadsheet Principles

GPS helps you to navigate through someone else's spreadsheet

Based on the European Spreadsheet Risk Interest Group (EuSpRIG) Guidance

http://www.eusprig.org/best-practice.htm



Good Practice is something of a MUST do

Each of the main principles of Good Practice Spreadsheet Modelling can be classified by MUST:

- **Maintainability**

  … which supports estimate or model **Inheritability** and implies **Simplicity and Accessibility** for future modification

- **Usability**

  … which infers that it is easy to use (**Simplicity** again) and **well-documented**

- **Security**

  … which guards against accidental corruption of the spreadsheet without denying **Accessibility** for future modification

- **Transparency**

  … which support Estimators' need for **Auditability** including **Verifiability & Validatability**. It also implies **Simplicity** and **Well-documented**

So, what 'MUST' Good Practice look like?

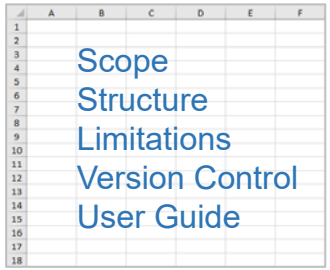
M	U	S	T
✓			✓

Supporting Documentation

GPS Principle:

The spreadsheet or model should be adequately documented in terms of its intended use, structure, limitations, user instructions, version control etc

Inherent Risk in failing to document adequately



- Users use the wrong version
- Users do not input all the data required
- Users expect the spreadsheet to calculate something outside of its true scope
- The Spreadsheet cannot easily be modified or maintained

Pragmatic Exceptions

- Spreadsheet is for personal, short-term use only and will be not be used as part of an audit trail or Basis of Estimate

Locked Cells and Worksheet Protection

M	U	S	T
		✓	

GPS Principle:

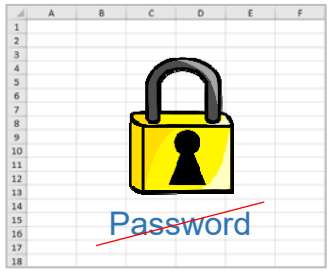
Only editable cells should be unlocked for the user to input. Worksheets and Workbook should be Protected from inadvertent editing of calculations and other non-editable cells. It is preferable not to use Password Protection.

Inherent Risk in not locking cells and protecting worksheets / workbooks

- Users might inadvertently overwrite a calculation creating an inappropriate result
- Password Protection leaves the spreadsheet or workbook open to “single point of failure”

GPS Recommendation

- Unlock input cells only. Apply Worksheet Protection without Passwords, allowing read only access to locked cells



Pragmatic Exceptions

- Spreadsheet is for personal, short-term use only and will be not be used by others

Hidden Worksheets, Rows and Columns

M	U	S	T
✓			✓

GPS Principle:

Worksheets, rows and columns should not be hidden from the user (including white on white etc)

Inherent Risk in hiding worksheets, rows or columns

- Users may inadvertently delete hidden elements by selecting around them and deleting or erasing the selection, causing errors or incorrect outcomes

GPS Recommendation

- Do not hide Worksheets, Rows or Columns from the user. Lock the appropriate cells and Protect the Worksheet

Pragmatic Exceptions

- Where users have left blank rows or columns to facilitate row and column alignment across worksheets, these might be hidden to improve the aesthetic look

M	U	S	T
		✓	✓

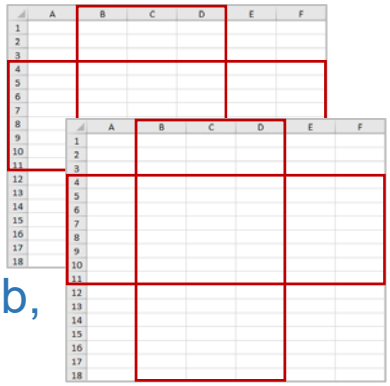
Column and Row Alignment

GPS Principle:

Related elements of data should be located in the same rows or columns within a single spreadsheet tab, or the same rows and columns across different tabs

Inherent Risk in not aligning rows and columns

- Errors are not detected because the logic is harder to follow (lateral alignment is easier to follow)
- Spreadsheet is more difficult to modify in the event of change



GPS Recommendation

- Leave blank rows and columns between sections on a single tab, and where there is no corresponding data

Pragmatic Exceptions

- Where spreadsheet logic is wholly contained in a visible or manageable range

M	U	S	T
	✓	✓	

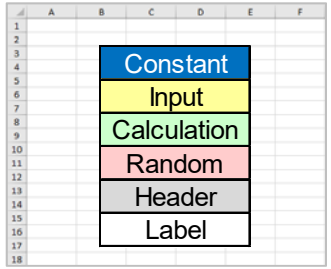
Consistent Colour Coding

GPS Principle:

Spreadsheets should use a consistent colour scheme throughout to denote the type of cell or tab. There should be a Colour Key on a separate Tab

Inherent Risk in not using consistent colour coding

- Users miss key inputs that have a material effect on the output because the need was not evident
- Users inadvertently over write a calculation and corrupt the output



GPS Recommendation

- Inputs, Constants, Calculations, Outputs and Special Features such as Monte Carlo Output should have a uniquely defined colour

Pragmatic Exceptions

- Spreadsheet is for personal, short-term use only and will be not be used by others

Hard-Coded Constants in Calculations

M	U	S	T
	✓	✓	

GPS Principle:

Do not use Hard-Coded Constants in Calculations e.g. **3.4***D14

Inherent Risk in using hard-coded constants in calculations

- When the constant becomes invalid, users fail to recognise that it needs to be changed, thus creating an error in the final outcome

GPS Recommendation

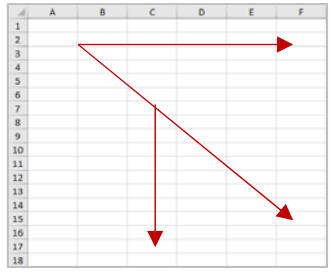
- Place constant in a separate cell that is “called” or referenced by the calculation

Pragmatic Exceptions

- Where the use of a constant is axiomatic and will never change later, such as using the power of 2 when something has to be squared, or 1 + an escalation rate
- Excel Function expects a parameter that is a constant e.g. MATCH(lookup_value,lookup_array,match_type) where match_type is -1, 0 or 1

Left to Right, Top to Bottom Readability

M	U	S	T
✓	✓		



GPS Principle:

A spreadsheet should read like a book, from left to right, top to bottom and front to back … as far as is reasonably possible,

Inherent Risk in not using left to right, top-bottom, front to back readability

- Future modifications may miss an important element of the logic chain if the logic flow is erratic, jumping forwards and then backwards

GPS Recommendation

- Calculations should reference cells to the left, above or on previous tabs

Pragmatic Exceptions

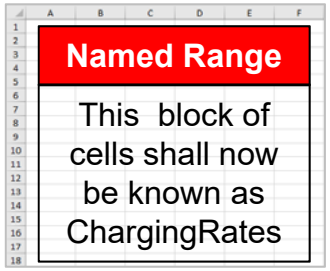
- Calculations that are dependent on the sum or average of a range of previous rows and columns can refer forward e.g. expressing a value as the percentage of a total, or the use of an interactive Dashboard

M	U	S	T
✓	✓		

Named Ranges

GPS Principle:

Use Named Ranges for frequently used cells or blocks of cells to enhance readability to simplify calculations



Inherent Risk in not using Named Ranges

- Verification and Validation takes longer and is more difficult to do thoroughly
- Future modifications can corrupt calculations

GPS Recommendation

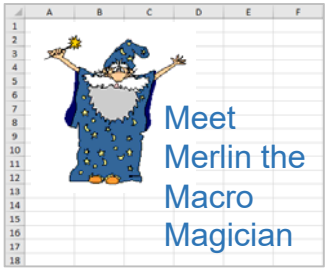
- Define Named Ranges in Excel and clearly mark these cells or ranges

Pragmatic Exceptions

- Where spreadsheet logic is wholly contained in a visible or manageable range

M	U	S	T
		✓	✓

Use of Macros



GPS Principle:

Macros that generate or modify data changes should be avoided

Inherent Risk in using macros

- Users change the structure of the workbook and the macros may read inappropriate cells, generating errors
- Macros are more difficult for user to follow and modify
- Independent Verification and Validation is delayed

GPS Recommendation

- Users can use Macros for Navigation

Pragmatic Exceptions

- Only use Macros that have Named Ranges for all cell and range references
- Define Macro-generated Output with a unique colour as a warning to users

Use of Array Formulae

M	U	S	T
✓	✓		✓

GPS Principle:

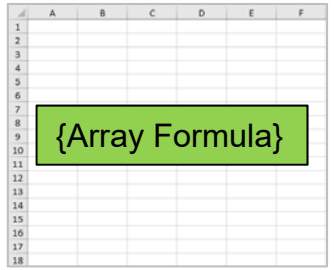
Array Formulae (where you have to press CTRL+SHIFT+ENTER to activate them) should be avoided wherever a simpler, practical alternative exists

Inherent Risk in using array formulae

- Many users do not understand Array Formulae and will not always spot an error
- Incorrectly modified, Array Formulae can create inappropriate, incorrect output but not always in a manner that is obvious

GPS Recommendation

- Look for standard Excel Functions that achieve the objective



Pragmatic Exceptions

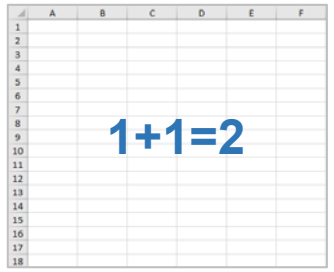
- Where they is no practical alternative to an Array Formulae, define a unique cell colour to highlight its existence to other users

M	U	S	T
✓			✓

Calculation Simplicity

GPS Principle:

Spreadsheet calculations should be as simple as possible



Inherent Risk in not keeping calculations simple

- Users cannot follow the calculation logic and avoid using the spreadsheet
- Complex calculations are more difficult to V&V and errors may still arise

GPS Recommendation

- Complex calculations (including multi-nested Condition Formulae) should be broken down into smaller, more transparent steps in adjacent columns or rows

Pragmatic Exceptions

- There are some (albeit only a few) calculations that require the use of an Array Formula. All Array Formulae class as complex calculations

Use Full Excel Function Syntax

M	U	S	T
✓			

GPS Principle:

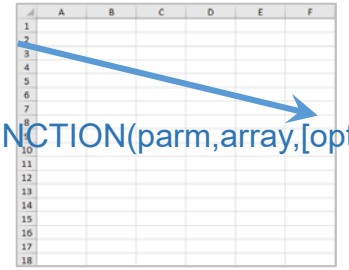
Always use the full Excel Syntax, even where Excel assumes a Default when a parameter is omitted

Inherent Risk in not using the full Excel Function Syntax

- Reliance on an Excel-defined default can be shown to produce unexpected or unintended results
- The spreadsheet logic may need to be modified in the future and the default may no longer be appropriate but the modification is missed

GPS Recommendation

- Always use the full Excel Syntax including the optional parameters



Pragmatic Exceptions

- No exceptions

M	U	S	T
		✓	

Active Links to External Spreadsheets

GPS Principle:

Active Links to External Spreadsheets that are used as Third Party Inputs, should be avoided where possible

Inherent Risk in using active links to external spreadsheets

- Third Party users change the feeder spreadsheet with consequential unexpected changes in our linked data that we are unable to explain
- Some Excel Functions do not work across Workbooks unless they are both open

GPS Recommendation

- Cut & Paste data from external spreadsheets as inputs.
- Have a dynamic link to a Checksum value to highlight when the 3rd Party data changes and is available to refresh by Cut & Paste

Pragmatic Exceptions

- If Direct Links have to be used for reasons of efficiency, set them with a specific cell colour and/or tab colour

M	U	S	T
	✓		

Unambiguous Units of Measure

GPS Principle:

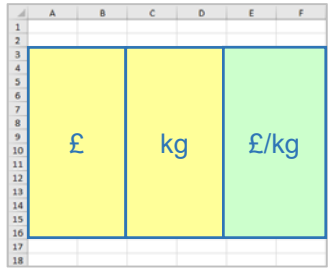
The units of measure or currency of a value in a cell should be unambiguous

Inherent Risk in not stating units of measure clearly

- Users may read the data as being absolute when it is scaled or vice versa e.g. £ instead of £k, or £ instead of $ or €
- Users may input data that is in different units of measure than was intended e.g. lbs instead of kg

GPS Recommendation

- Place the Unit of Measure / Currency in the Column Header, or the Row Label, or use Cell formatting to include the unit, or in an adjacent blank cell



Pragmatic Exceptions

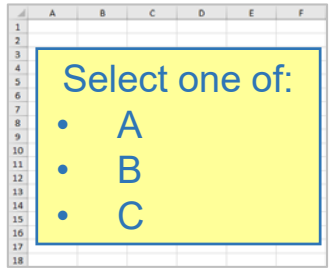
- No exceptions

M	U	S	T
	✓		

Input Data Validation

GPS Principle:

Use Excel's Data Validation feature to control inputs by users



Inherent Risk in not using data validation

- Users input inappropriate values outside of the intended or expected range creating potentially invalid outputs

GPS Recommendation

- Prevent user inputting values outside or a Min/Max range, or non-integer when an integer is expected using Excel's Data Validation feature
- Create lists of allowable text input values in conjunction with Named Ranges e.g. Manufacturing but not Production

Pragmatic Exceptions

- Spreadsheet is for personal, short-term use only and will be not be used by others

Independent Verification and Validation

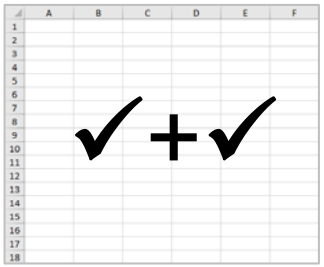
M	U	S	T
		✓	

GPS Principle:

Spreadsheets should be independently verified and validated against the defined business objective

Inherent Risk in not having spreadsheets independently V&V'd

- Spreadsheet developers may see what they expect to see and miss logic errors, typing errors and other input errors



GPS Recommendation

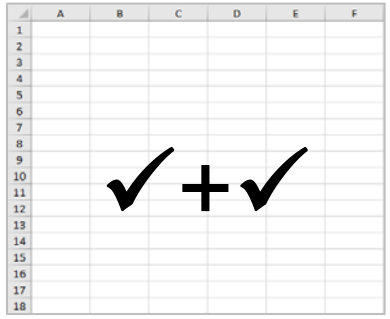
- All Business Critical spreadsheets are subject to independent Verification and Validation by a suitably trained auditor

Pragmatic Exceptions

- Spreadsheet is for personal, short-term use only and will be not be used by others

Verification and Validation

Not just about detectig errors



Verification and Validation?

What's the difference? Which is which?

Are the Inputs appropriate?

- **Validation** focuses on whether the assumptions and data used in a spreadsheet or model are accuracy and appropriate for their intended purpose
- **Verification** focuses on whether the calculations and logic of a spreadsheet or model are accurate and appropriate for their intended purpose

Does it calculate correctly?

It is not just Pass or Fail

- If the V&V assessment detects an error, then the Spreadsheet is usually "failed" or sent back for correction before use
- Just because V&V has not detected any errors, it does not mean that the spreadsheet is risk free (or even error free)
- A good V&V Report will identify risks in the ongoing use of a spreadsheet, not just current errors
- Each organisation needs to decide how much risk is acceptable

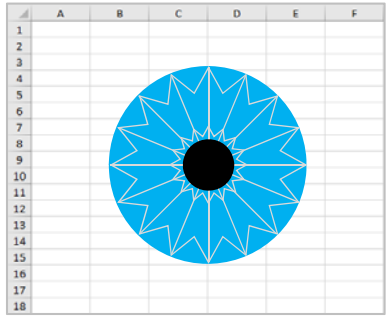
Agree or Disagree?

These "principles" are not necessarily absolute, or even complete

- Each organisation should evaluate the risks of current spreadsheet practice and decide on what is important to them
  - How they can better mitigate these risks with improved and consistent spreadsheet discipline?
  - Encapsulate their expectations in a Policy Statement
- Here, we have discussed 16 GPS Principles, but different organisations can tailor the list to their needs and circumstances
- Having created a "policy" on acceptable spreadsheet practice, organisations might want to consider tagging each business critical spreadsheet with an *IRiS* Rating …
- So, back to the original question, what is *IRiS* ? … and how can *IRiS* help?

IRiS Scorecard

Keeping an eye on those risks



IRiS Scorecard

- We can use the following as a self-assessment tool when we have created an important spreadsheet or model

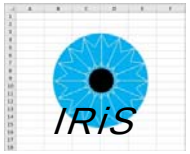
  Or, it can be used by someone performing a Validation and Verification of someone else's spreadsheet

- Against each of the GPS Modelling Principles, we can score them as follows:

	Risk Score
• Follows Good Practice principles well	0
• Occasionally fails to follow Good Practice principles	1
• Several instances where Good Practice is not observed	2
• Multiple instances of failing Good Practice principles	3

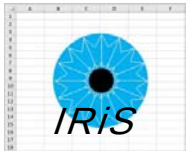
- By taking an average we get an overall view of the Inherent Risk in our (or someone else's) Spreadsheet
  - Alternatively, to increase the weighting of higher risk scores, we can take the Square Root of the Average Squares of the Scores

IRiS Scorecard 1 of 5



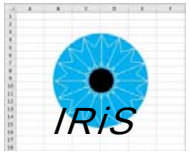
Good Practice Feature	Risk Score			
	0 - Low Risk	1 - Medium Risk	2 - High Risk	3 – Very High Risk
Level of Documentation	Full documentation available	Most documentation Available	Little documentation available	No documentation available
Hard-Coded Constants in Calculations	No hard-coded constants in calculation cells except axiomatic formula values	Some instances of hard-coded constants in calculation cells of small spreadsheet	Some instances of hard-coded constants in calculation cells of large spreadsheet	Multiple instances of hard-coded constants in calculation cells
Column and Row Alignment	All corresponding columns and rows align across multiple worksheets	Some corresponding rows do not align across multiple columns of a single worksheet	Some instances where corresponding columns and rows are not aligned across multiple worksheets	Several instances where corresponding columns and rows are not aligned across multiple worksheets
Full Excel Syntax	Full syntax is being used for all Excel functions	Occasional instances of syntax being abbreviated for a small number of Excel functions	Multiple instances of syntax consistently being abbreviated for a range of Excel Functions	Multiple but inconsistent instances of syntax being abbreviated for a range of Excel Functions

IRiS Scorecard 2 of 5



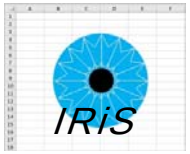
Good Practice Feature	Risk Score			
	0 - Low Risk	1 - Medium Risk	2 - High Risk	3 – Very High Risk
Named Ranges	Extensive use of Named Ranges, proportionate to spreadsheet size and complexity	Some use of Named Ranges, proportionate to spreadsheet size and complexity	Small simple spreadsheet with no named ranges	Large complex spreadsheet with no named ranges
Input Data Validation	Extensive use made of Input Data Validation with User Help explanations	Wide use made of Input Data Validation with User Help explanations	Limited use of Input Data Validation	No use of Input Data Validation on a complex spreadsheet with multiple input variables
Left to Right Top to Bottom Readability	Consistent flow of workbook logic from left to right, top to bottom, and front to back	Generally good adherence to workbook logic flowing from left to right, top to bottom, and front to back	Logic flow is inconsistent with a number of instances of jumping from right to left, or bottom to top, or back to front	Logic flow is haphazard jumping frequently from right to left, or bottom to top, or back to front

IRiS Scorecard 3 of 5



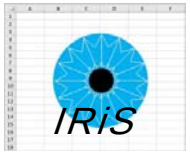
Good Practice Feature	Risk Score			
	0 - Low Risk	1 - Medium Risk	2 - High Risk	3 – Very High Risk
Clarity of Units of Measure	All cells, columns, rows and worksheet tabs are clearly labelled with their units of measurement / scales	Most cells, columns, rows and worksheet tabs are labelled with their units of measurement / scales	Basic cell value measurement units are visible but scale units (k, m, b) are frequently missing	Cell value measurement unit scales are largely missing
Consistent Colour Coding	Organisation approved standard colours are used consistently and correctly	Generally, cell colours are not used, or there are some inconsistencies in the application of standard colours	Non-standard colours are used instead of organisation approved ones	Similar cell types are coloured differently
Locked Cells, Worksheet and Workbook Protection	Input cells are unlocked, other cells are locked, worksheets and workbooks are protected but without password	Input cells are unlocked, other cells are locked, worksheets are password protected or the workbook is unprotected	Some worksheets or the workbook are unprotected leaving those cells unlocked and vulnerable to change	All worksheets and workbook are unprotected leaving all cells unlocked and vulnerable to change or deletion

IRiS Scorecard 4 of 5



Good Practice Feature	Risk Score			
	0 - Low Risk	1 - Medium Risk	2 - High Risk	3 – Very High Risk
Active External Links	There are no active links to external data sources (other than checksum values)	There are some active links that simply read external data sources	There are some active links that perform calculations on external data sources	There are multiple links to or high value links calculated from external data sources
Use of Macros	Any macros are used for navigation only	Macros are used to calculate cell values, but they all use Named Ranges rather than absolute cell references	Some macros are used to generate cell values, and logic uses absolute cell references rather than Named Ranges	Multiple macros are used to generate cell values, and logic uses absolute cell references rather than Named Ranges
Independent Model Verification & Validation	Independently V&V'ed by trained Assessor	Simple Peer review (Dip Check)	Multiple Checksum Vets (Self-inspection only)	No checking performed
Use of Array Formula	No Array Formulae are used	One Array Formula is being used but explanatory text is available	One Array Formula is being used without documented notes	Multiple instances of Array Formulae being used

IRiS Scorecard 5 of 5



Good Practice Feature	Risk Score			
	0 - Low Risk	1 - Medium Risk	2 - High Risk	3 – Very High Risk
Hidden Worksheets, Rows and Columns	No hidden worksheets, columns or rows	Worksheet hidden but no hidden rows or columns	Some rows and columns are hidden	Multiple instances of hidden rows and columns
Simplicity of Calculations	All calculations are simple	Calculations are complex but documented	Multiple nested condition formulae are being used without documentation	Multiple nested condition formulae are being used without documentation

This list can be modified to the needs of individual organisation

If you do use or modify ***IRiS***, please have the courtesy of citing this presentation as the source

IRiS Scorecard Example

Feature	Risk Score				
	0 - Low Risk	1 - Below Average Risk	2 - Above Average Risk	3 - High Risk	Score
Level of Documentation	Full documentation available	Most documentation Available	Little documentation available	No documentation available	2
Hard-Coded Constants in Calculations	No hard-coded constants in calculation cells except axiomatic formula values	Some instances of hard-coded constants in calculation cells of small spreadsheet	Some instances of hard-coded constants in calculation cells of large spreadsheet	Multiple instances of hard-coded constants in calculation cells	1
Column and Row Alignment	All corresponding columns and rows align across multiple worksheets	Some corresponding rows do not align across multiple columns of a single worksheet	Some instances where corresponding columns and rows are not aligned across multiple worksheets	Several instances where corresponding columns and rows are not aligned across multiple worksheets	0
Full Excel Syntax	Full syntax is being used for all Excel functions	Occasional instances of syntax being abbreviated for a small number of Excel functions	Multiple instances of syntax consistently being abbreviated for a range of Excel Functions	Multiple but inconsistent instances of syntax being abbreviated for a range of Excel Functions	0
Named Ranges	Extensive use of Named Ranges, proportionate to spreadsheet size and complexity	Some use of Named Ranges, proportionate to spreadsheet size and complexity	Small simple spreadsheet with no named ranges	Large complex spreadsheet with no named ranges	1
Input Data Validation	Extensive use made of Input Data Validation with User Help explanations	Wide use made of Input Data Validation with User Help explanations	Limited use of Input Data Validation	No use of Input Data Validation on a complex spreadsheet with multiple input variables	1
Left to Right Top to Bottom Readability	Consistent flow of workbook logic from left to right, top to bottom, and front to back	Generally good adherence to workbook logic flowing from left to right, top to bottom, and front to back	Logic flow is inconsistent with a number of instances of jumping from right to left, or bottom to top, or back to front	Logic flow is haphazard jumping frequently from right to left, or bottom to top, or back to front	0
Clarity of Units of Measure	All cells, columns, rows and worksheet tabs are clearly labelled with their units of measurement / scales	Most cells, columns, rows and worksheet tabs are labelled with their units of measurement / scales	Basic cell value measurement units are visible but scale units (k, m, b) are frequently missing	Cell value measurement unit scales are largely missing	0
Consistent Colour Coding	Organisation approved standard colours are used consistently and correctly	Generally, cell colours are not used, or there are some inconsistencies in the application of standard colours	Non-standard colours are used instead of organisation approved ones	Similar cell types are coloured differently	1
Locked Cells, Worksheet Protection	Input cells are unlocked, other cells are locked, worksheets and workbooks are protected but without password	Input cells are unlocked, other cells are locked, worksheets are password protected or the workbook is unprotected	Some worksheets or the workbook are unprotected leaving those cells unlocked and vulnerable to change	All worksheets and workbook are unprotected leaving all cells unlocked and vulnerable to change or deletion	3
Active External Links	There are no active links to external data sources (other than checksum values)	There are some active links that simply read external data sources	There are some active links that perform calculations on external data sources	There are multiple links to or high value links calculated from external data sources	1
Use of Macros	Any macros are used for navigation only	Macros are used to calculate cell values, but they all use Named Ranges rather than absolute cell references	Some macros are used to generate cell values, and logic uses absolute cell references rather than Named Ranges	Multiple macros are used to generate cell values, and logic uses absolute cell references rather than Named Ranges	0
Independent Model Verification & Validation	Independently V&V'ed by trained Assessor	Simple Peer review (Dip Check)	Multiple Checksum Vets (Self-inspection only)	No checking performed	1
Use of Array Formula	No Array Formulae are used	One Array Formula is being used but explanatory text is available	One Array Formula is being used without documented notes	Multiple instances of Array Formulae being used	0
Hidden Worksheets, Rows, Columns	No hidden worksheets, columns or rows	Worksheet hidden but no hidden rows or columns	Some rows and columns are hidden	Multiple instances of hidden rows and columns	0
Simplicity of Calculations	All calculations are simple	Calculations are complex but documented	Multiple nested condition formulae are being used without documentation	Multiple nested condition formulae are being used without documentation	2
					0.8

Square Root Average Square Score = 1.2

Score the spreadsheet against each GPS Principle

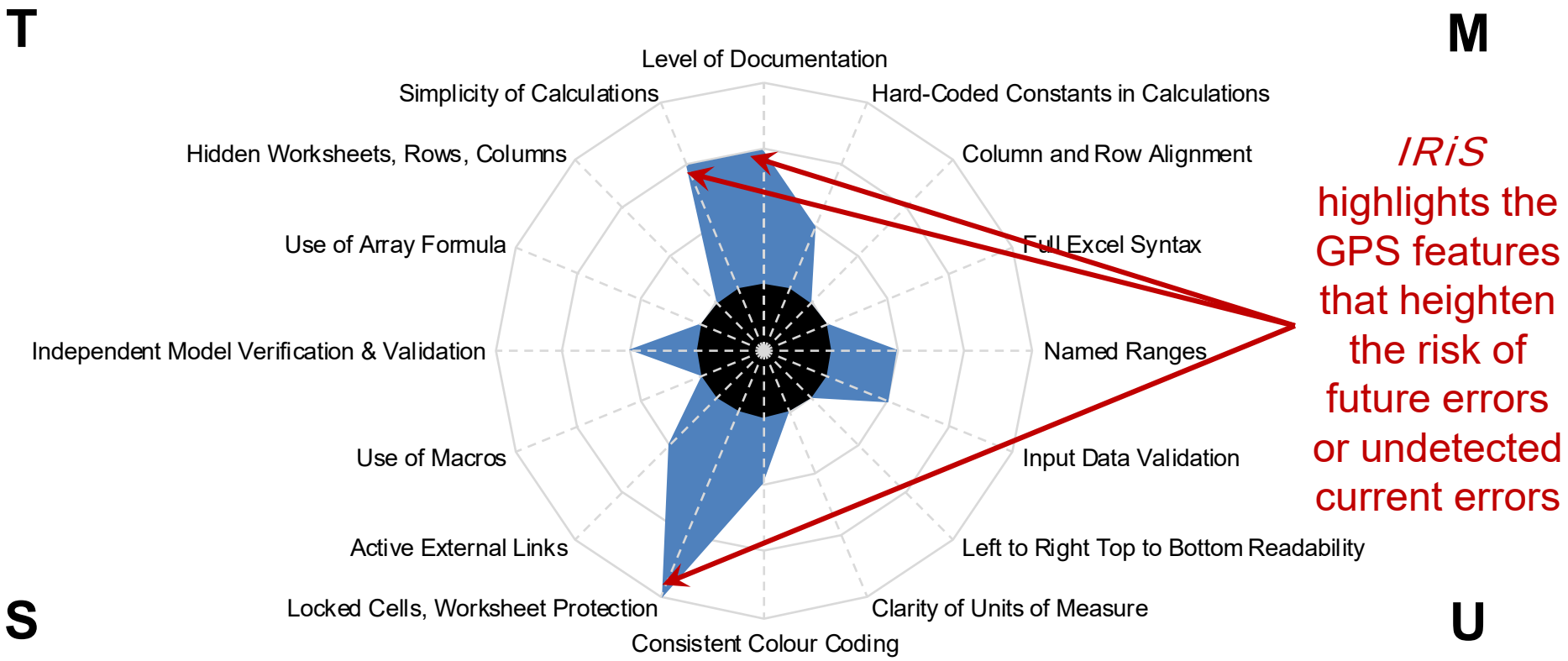
Or those agreed in line with your organisation's policy

Look at whether Scores of 2 or 3 can be addressed easily to return them to 0 or 1 risks

Take a simple average of the scores across the agreed principles

IRiS Chart

Inherent Risk in Spreadsheets Score



IRiS highlights the GPS features that heighten the risk of future errors or undetected current errors

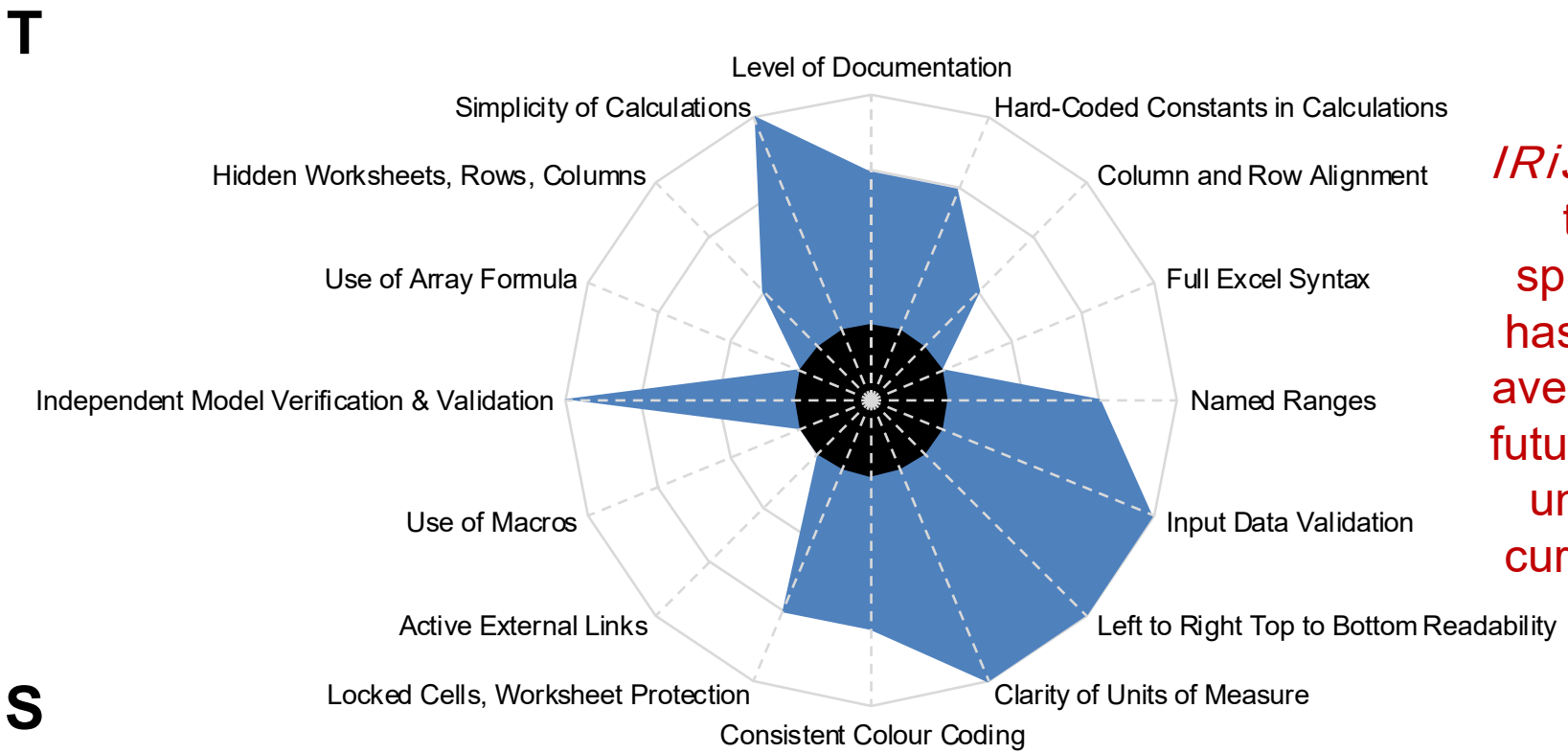
Source: Author

Average Score = 0.8

Square Root Average Square Score = 1.2

IRiS Chart (Further Example)

Inherent Risk in Spreadsheets Score



IRiS highlights that this spreadsheet has an above average risk of future errors or undetected current errors

Source: Author

Average Score 1.7

Square Root Average Square Score = 2.05

IRiS Scorecard Example

- The more we see of the *IRiS*, then the more we need to keep an eye on the Inherent Risk in our Spreadsheets

- Follow GPS Principles (or define those that work for your organisation)
- Bearing in mind that:
  - GPS are a guide to reducing undetected errors
  - No spreadsheet is perfect
  - Even the GPS Principles can conflict with each other; it's about achieving a pragmatic balance

- Let's not turn a blind eye to the Inherent Risk in Spreadsheets
- *IRiS* can help you visualise the risks and focus on mitigating them

Thank you for Listening

Any Questions?

Any Comments or Suggestions?