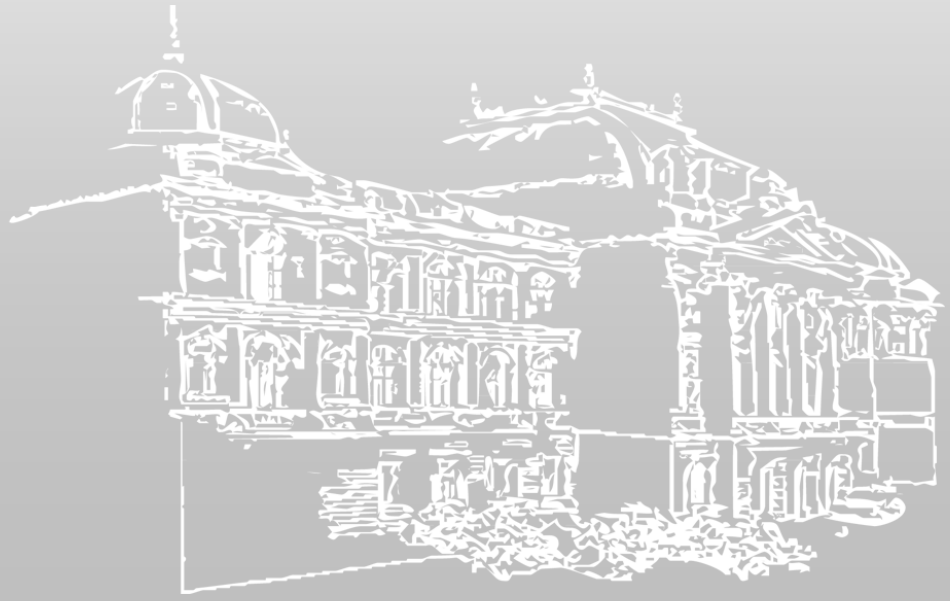
Fault localization in the light of faulty user input

Birgit Hofer, and Franz Wotawa
Graz University of Technology, Austria



2015 IEEE Int. Conference on Software Quality, Reliability & Security (QRS), Vancouver, Canada, August 2015

Why spreadsheets?

- Range
  - Used in nearly every company

- Fields of application
  - Forecasting
  - Financial Reporting

  →Basis for decisions

- Can be seen as programming language

  → Reuse ideas for other programming languages

To err is human.

3-5 % error rate when writing formulas*

* Source: Panko "*Thinking is bad: Implications of human error research for spreadsheet research and practice*," CoRR, 2008.

Outline

1. Running Example
2. Spectrum-based Fault Localization
3. Misclassifications
4. Impact Analysis
   a. Theoretical Analysis
   b. Empirical Evaluation
5. Conclusion

Running Example

	A	B	C	D	E	F	G	H	I	J
1	**Worker**	**Mo**	**Tu**	**We**	**Th**	**Fr**	**Total**	**Overtime**	**$/h**	**Total wage**
2	**Jones**	9	9	8	8	8	42	2	20	860
3	**Smith**	9	9	10	7	8	43	1	24	1044
4	**Total**						85	3		
5										
6	**Regular working time**				40					
7	**Working hours Jones**				42					
8	**Working hours Smith**				43					

My wage is too low!

G	H	I	J
Total	**Overtime**	**$/h**	**Total wage**
=SUM(B2:F2)	=IF(G2>E$6;G2-E6;0)	20	=I2*(G2+H2*0,5)
=SUM(B3:F3)	=IF(G3>E$6;G3-E7;0)	24	=I3*(G3+H3*0,5)
=SUM(G2:G3)	=SUM(H2:H3)		

Birgit Hofer and Franz Wotawa: „Fault localization in the light of faulty user input “
QRS, Vancouver, Canada, August 2015

Running Example

	A	B	C	D	E	F	G	H	I	J
1	**Worker**	**Mo**	**Tu**	**We**	**Th**	**Fr**	**Total**	**Overtime**	**$/h**	**Total wage**
2	**Jones**	9	9	8	8	8	=SUM(B2:F2)	=IF(G2>E$6;G2-E6;0)	20	=I2*(G2+H2*0,5)
3	**Smith**	9	9	10	7	8	=SUM(B3:F3)	=IF(G3>E$6;G3-E7;0)	24	=I3*(G3+H3*0,5)
4	**Total**						=SUM(G2:G3)	=SUM(H2:H3)		
5										
6	**Regular working time**				40					
7	**Working hours Jones**				=G2					
8	**Working hours Smith**				=G3					

$$\mathrm{CONE}(c) = c \cup \bigcup_{c' \in \rho(c)} \mathrm{CONE}(c')$$

CONE(J3) = {E7, G2, G3, H3, J3}

CONE(H4) = {E7, G2, G3, H2, H3, H4}

CONE(G4) = {G2, G3, G4}

CONE(J2) = {E7, G2, G3, H3, J3}

Spectrum-based Fault Localization with the Ochiai Similarity Coefficient

$$\mathrm{SC}_O(c) = \frac{a_{11}(c)}{\sqrt{(a_{11}(c) + a_{10}(c))(a_{11}(c) + a_{01}(c))}}$$

$$a_{11}(c) = |\{(c', d) \in O | c \in \textsc{Cone}(c') \wedge d = \text{false}\}|$$

$$a_{10}(c) = |\{(c', d) \in O | c \in \textsc{Cone}(c') \wedge d = \text{true}\}|$$

$$a_{01}(c) = |\{(c', d) \in O | c \notin \textsc{Cone}(c') \wedge d = \text{false}\}|$$

$$\Rightarrow \quad a_{11}, a_{10}, a_{01} \geq 0$$

IF ($a_{11}(c) + a_{10}(c) = 0$ OR $a_{11}(c) + a_{01}(c) = 0$)
then $\mathrm{SC}_O(c) = 0$

Birgit Hofer and Franz Wotawa: „Fault localization in the light of faulty user input “
QRS, Vancouver, Canada, August 2015

Running Example

Cell	G4	H4	J2	J3
E7		●		●
E8				
G2	●	●	●	●
G3	●	●		●
G4	●			
H2		●	●	
H3		●		●
H4		●		
J2			●	
J3				●
Classification d	**true**	**false**	**true**	**false**

Running Example

Cell	G4	H4	J2	J3	SC	Rank
E7		●		●	1.000	1
E8					0	-
G2	●	●	●	●	0.707	4
G3	●	●		●	0.816	3
G4	●				0	-
H2		●	●		0.500	7
H3		●		●	1.000	1
H4		●			0.707	4
J2			●		0	-
J3				●	0.707	4
Classification d	**true**	**false**	**true**	**false**		

Birgit Hofer and Franz Wotawa: „Fault localization in the light of faulty user input “
QRS, Vancouver, Canada, August 2015

To err is human.

Assuming a 3-5 % error rate

26-40 %

chance to make at least 1 mistake
when classifying 10 cell values.

Birgit Hofer and Franz Wotawa: „Fault localization in the light of faulty user input “
QRS, Vancouver, Canada, August 2015

Never assume **perfect** user knowledge when **evaluating** fault localization techniques for **end-user** software.

What are the impacts of a misclassification?

Outline

1. Running Example
2. Spectrum-based Fault Localization
3. Misclassifications
4. Impact Analysis
   - a. Theoretical Analysis
   - b. Empirical Evaluation
5. Conclusion

Theoretical analysis of the impact of a misclassification on the similarity coefficient

	Misclassification of cell c as	
	erroneous	correct
$c_1 \in CONE(c)$	increase	
$c_2 \notin CONE(c)$		

Premises:

$a_{11} + a_{10} > 0$

$a_{11} + a_{01} > 0$

$$SC_O^{\#}(c_1) = \frac{a_{11}+1}{\sqrt{(a_{11}+1+a_{10}-1)(a_{11}+1+a_{01})}}$$

$$= \frac{a_{11}+1}{\sqrt{(a_{11}+a_{10})(a_{11}+a_{01}+1)}}.$$

$$SC_O^{\#}(c_1) > SC_O^{*}(c_1)$$

Theoretical analysis of the impact of a misclassification on the similarity coefficient

	Misclassification of cell c as	
	erroneous	correct
$c_1 \in CONE(c)$	a_{11}++ a_{10}−−	a_{11}−− a_{10}++
$c_2 \notin CONE(c)$	a_{01}++	a_{01}−−

Theoretical analysis of the impact of a misclassification on the similarity coefficient

	Misclassification of cell c as	
	erroneous	correct
$c_1 \in CONE(c)$	increase	decrease or equal
$c_2 \notin CONE(c)$	decrease*	increase**

*Special case $a_{11} + a_{10} = 0 \vee a_{11} + a_{01} = 0$:
$SC^*_O(c_2) = SC^\#_O(c_2) = 0$

**Special case $a_{11} + a_{10} = 0 \vee a_{11} + a_{01} = 1$:
$SC^*_O(c_2) = SC^\#_O(c_2) = 0$

Influence of Misclassification in Running Example

Cell	G4	H4	J2	J3	SC	Rank
E7		●		●	0.816	3
E8					0	-
G2	●	●	●	●	0.866	2
G3	●	●		●	1.000	1
G4	●				0.577	5
H2		●	●		0.408	8
→ H3		●		●	0.816	3
H4		●			0.577	5
J2			●		0	-
J3				●	0.577	5
Classification d	**false**	**false**	**true**	**false**		

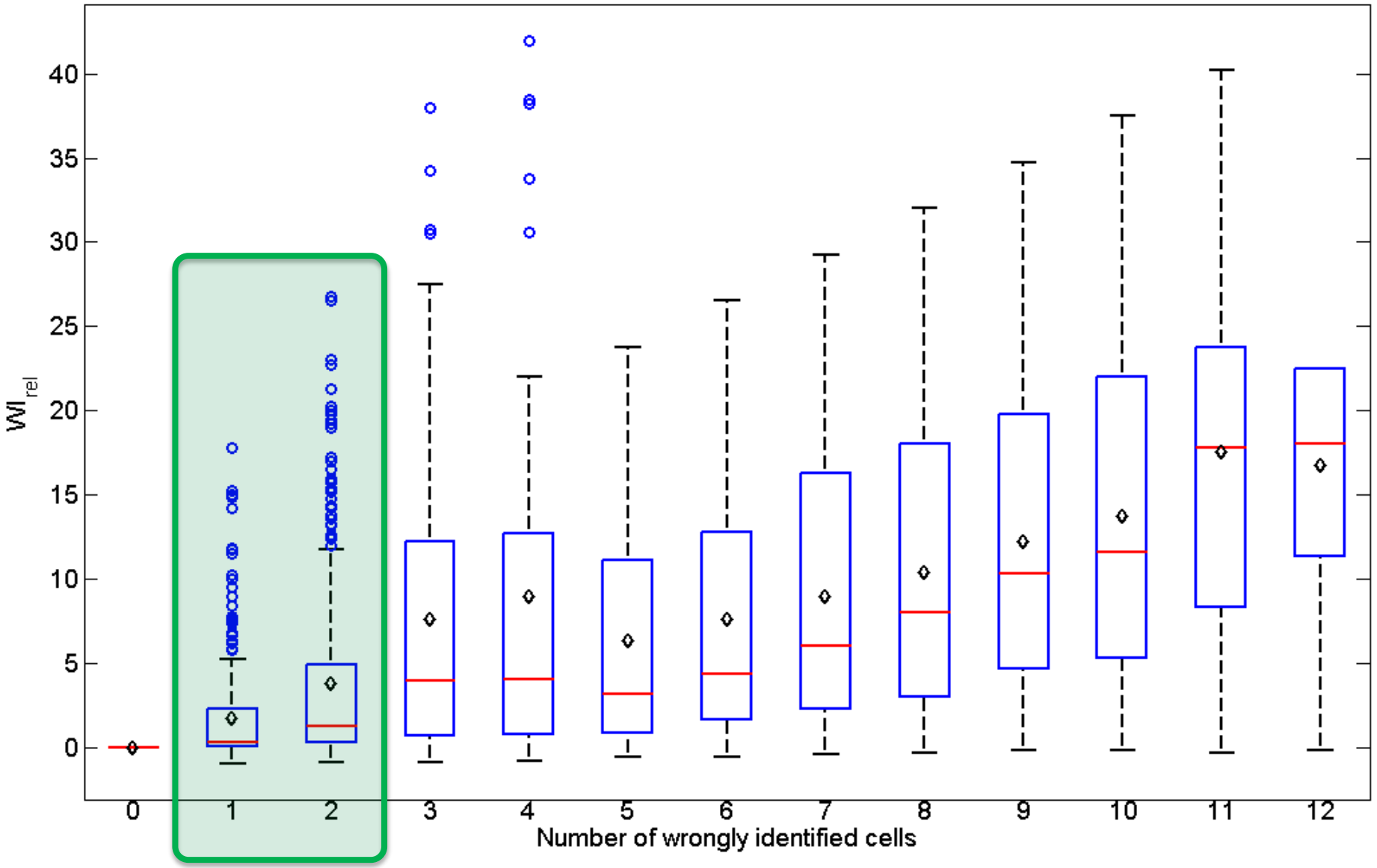
Empirical evaluation

- Integer spreadsheet corpus
  - 218 faulty spreadsheets
- Relative Worsening/Improvement

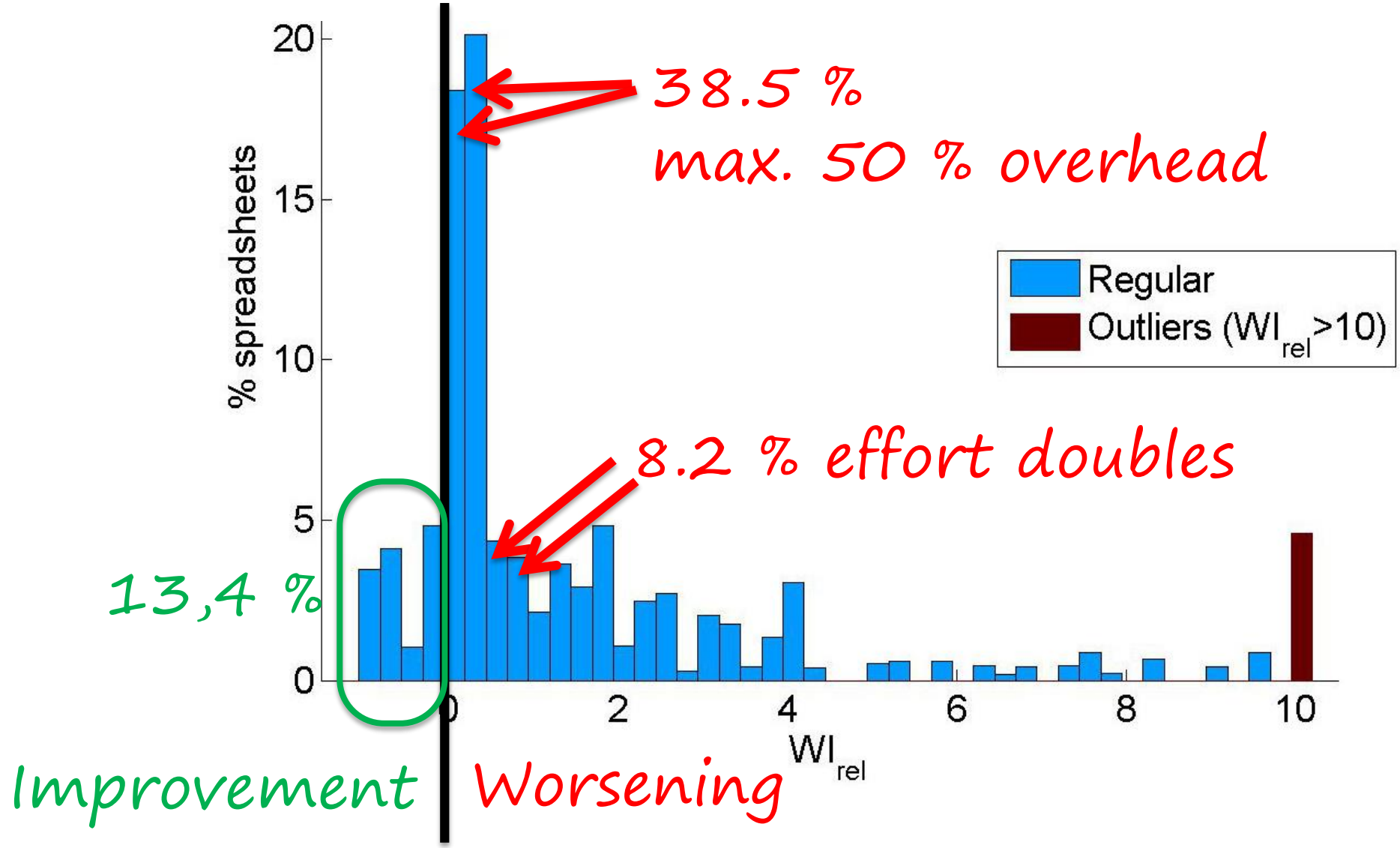
$$\mathrm{WI}_{\mathrm{rel}} = \frac{R_{\mathrm{avg}}}{R^*_{\mathrm{avg}}} - 1$$

$$R_{\mathrm{avg}} = |\{i | \mathrm{SC}_O(i) > \mathrm{SC}_O(f)\}| + \\ + \frac{|\{j | \mathrm{SC}_O(j) = \mathrm{SC}_O(f)\}|}{2} + 0.5$$

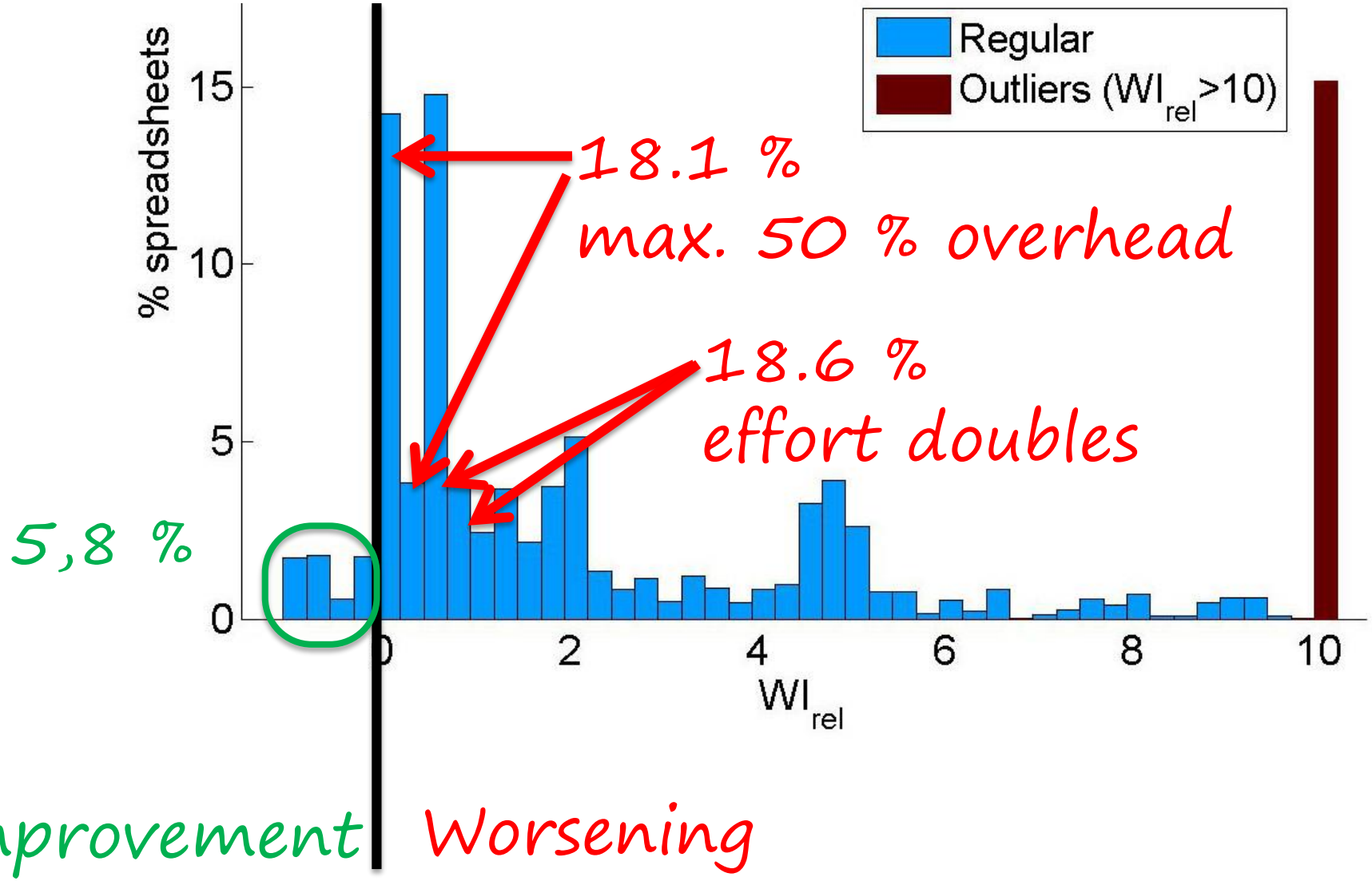
Relative Worsening/Improvement



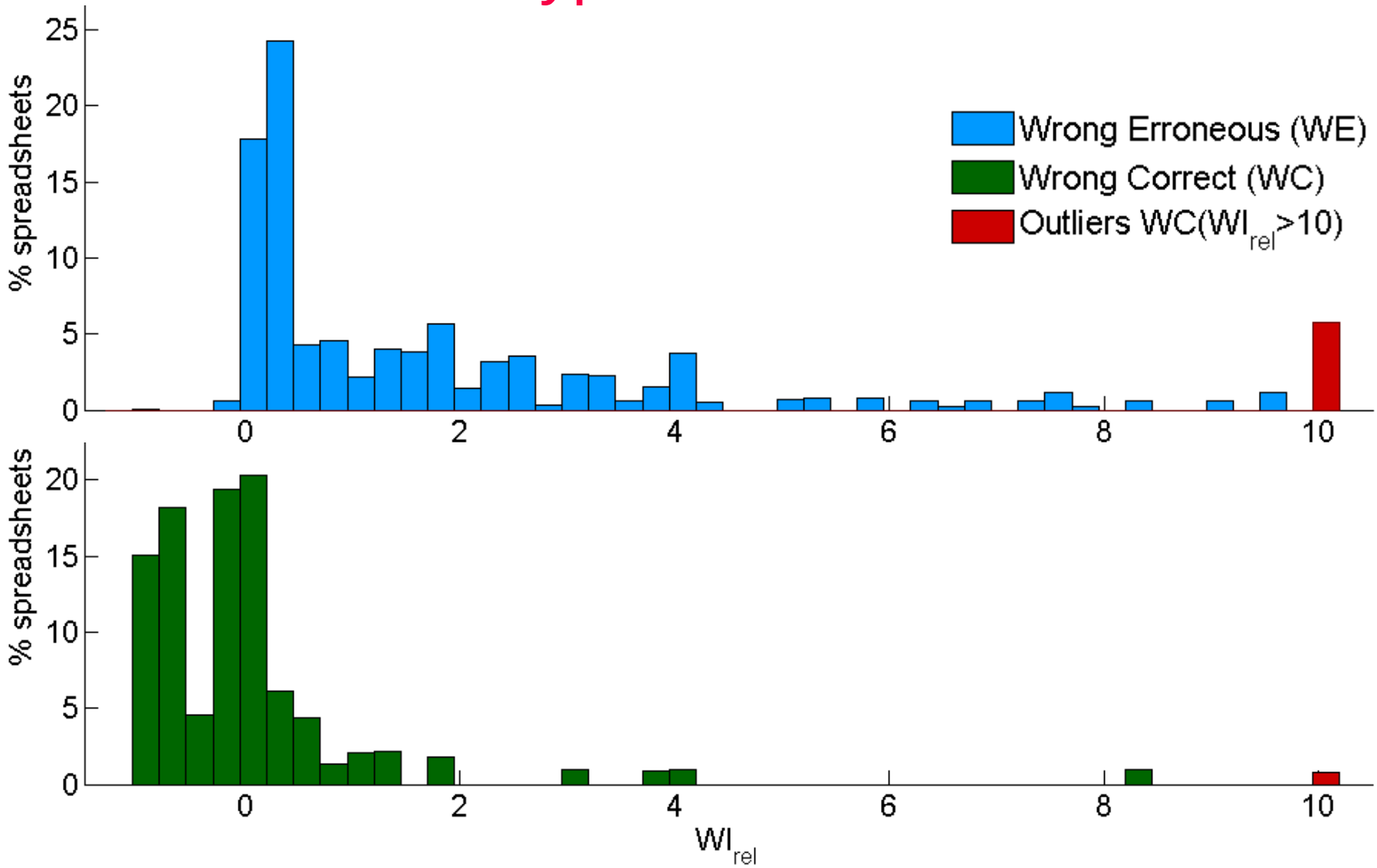
Effect histogram for 1 misclassification



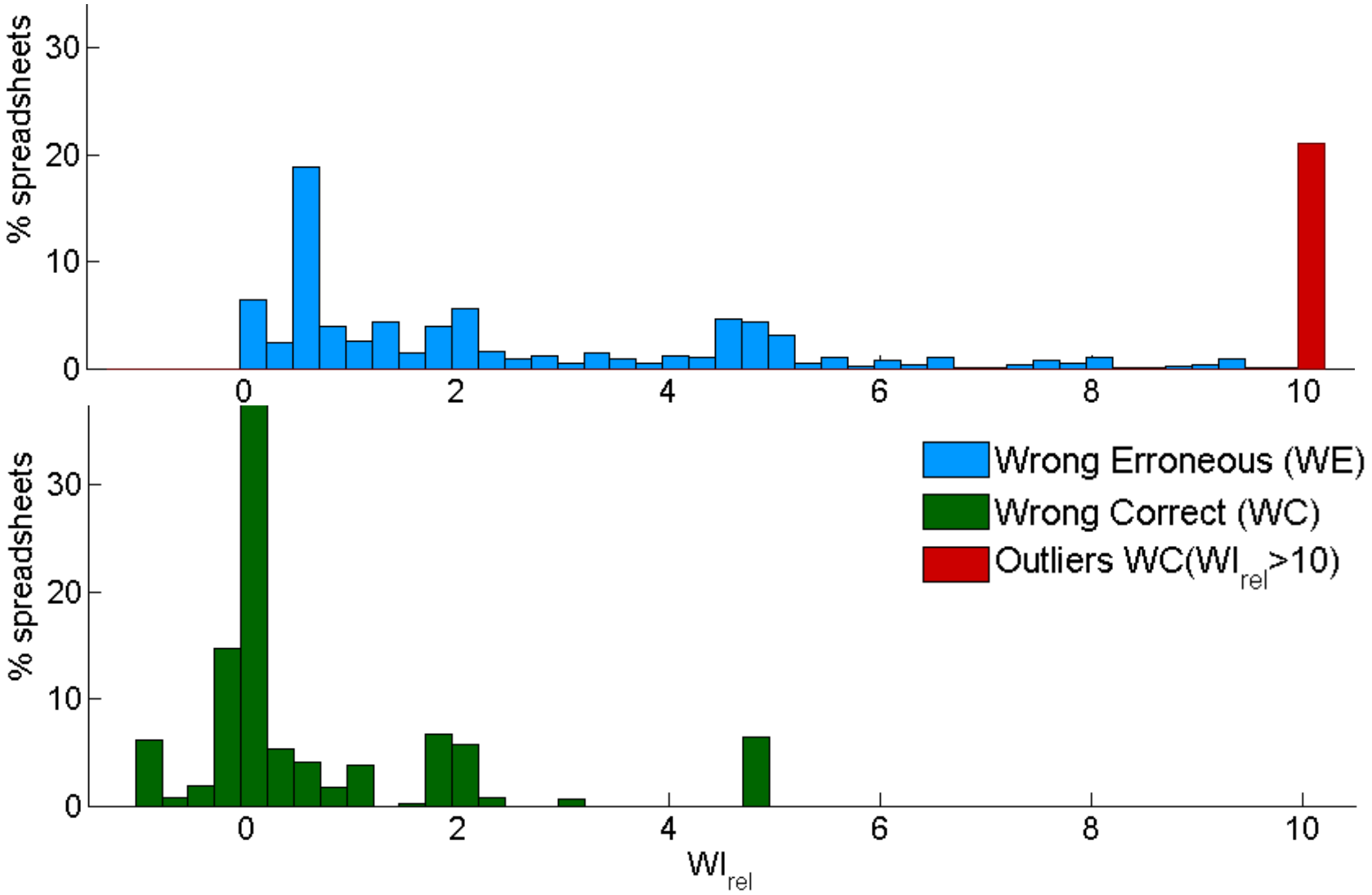
Effect histogram for 2 misclassifications



Effect histogram for 1 misclassification for different types of misclassifications



Effect histogram for 2 misclassifications for different types of misclassifications



To err is human.

3-5 % error rate when writing formulas*

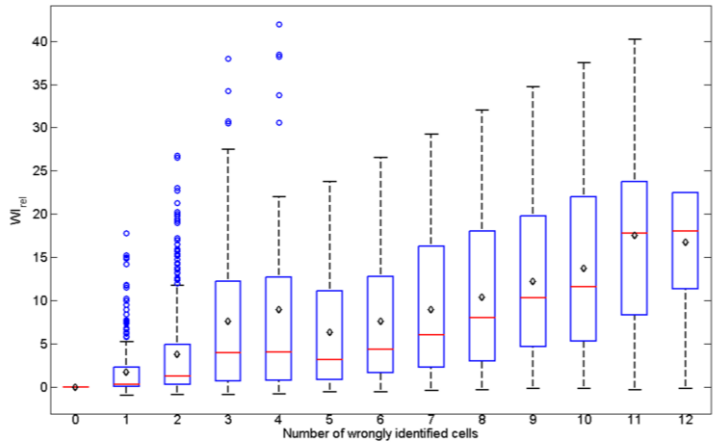
Theoretical analysis of the impact of a misclassification on the similarity coefficient

	Misclassification of cell c as	
	erroneous	correct
$c_1 \in CONE(c)$	increase	decrease or equal
$c_2 \notin CONE(c)$	decrease*	increase**

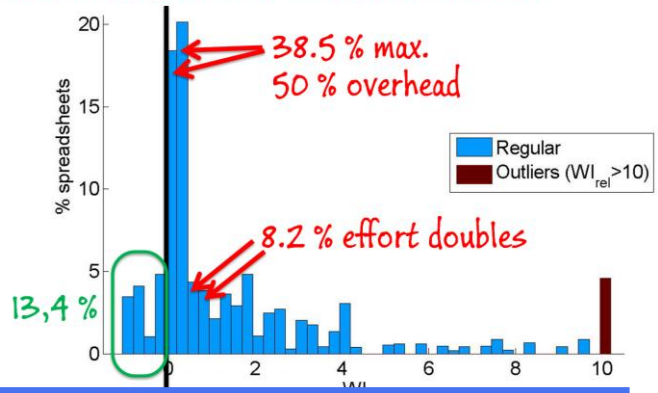
$= 0$:

$= 1$:

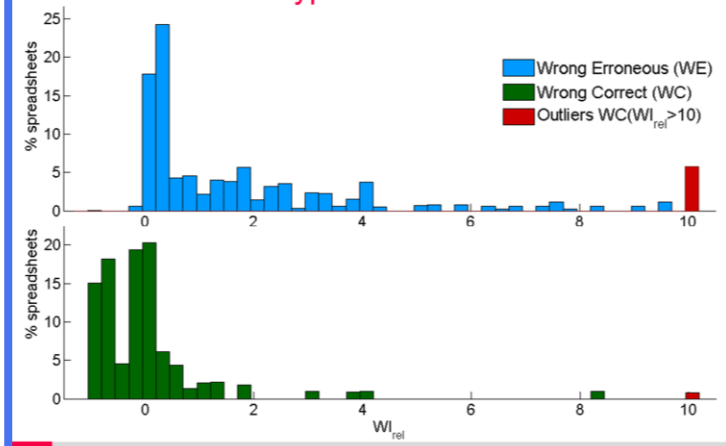
Relative Worsening/Improvement



Effect histogram for 1 misclassification



Effect histogram for 1 misclassification for different types of misclassifications



Thank you!
Questions?