S C I E N C E ▪ P A S S I O N ▪ T E C H N O L O G Y

Qualitative deviation models vs. quantitative models for fault localization in spreadsheets

Birgit Hofer, Iulia Nica, and Franz Wotawa

Graz University of Technology

FWF Der Wissenschaftsfonds.

DFG Deutsche Forschungsgemeinschaft

Funded by the Austrian Science Fund (FWF, contract I2144) and the Deutsche Forschungsgemeinschaft (DFG,contract JA 2095/4-1)

Spreadsheet Errors

JPMorganChase

~ 300 million EUR damage

Spreadsheet Errors

Paper published by Reinhart & Rogoff

Many governments used it for decisions

Immense loss of reputation

	A	B	I	J	K	L	M
2			Real GDP growth				
3			Debt/GDP				
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3,7	3,0	3,5	1,7	5,5
27	Minimum		1,6	0,3	1,3	-1,8	0,8
28	Maximum		5,4	4,9	10,2	3,6	13,3
29							
30	US	1946-2009	n.a.	3,4	3,3	-2,0	n.a.
31	UK	1946-2009	n.a.	2,4	2,5	2,4	n.a.
32	Sweden	1946-2009	3,6	2,9	2,7	n.a.	6,3
33	Spain	1946-2009	1,5	3,4	4,2	n.a.	9,9
34	Portugal	1952-2009	4,8	2,5	0,3	n.a.	7,9
35	New Zealand	1948-2009	2,5	2,9	3,9	-7,9	2,6
36	Netherlands	1956-2009	4,1	2,7	1,1	n.a.	6,4
37	Norway	1947-2009	3,4	5,1	n.a.	n.a.	5,4
38	Japan	1946-2009	7,0	4,0	1,0	0,7	7,0
39	Italy	1951-2009	5,4	2,1	1,8	1,0	5,6
40	Ireland	1948-2009	4,4	4,5	4,0	2,4	2,9
41	Greece	1970-2009	4,0	0,3	2,7	2,9	13,3
42	Germany	1946-2009	3,9	0,9	n.a.	n.a.	3,2
43	France	1946-2022	4,9	2,7	3,0	n.a.	5,2
44	Finland	1946-2023	3,8	2,4	5,5	n.a.	7,0
45	Denmark	1946-2024	3,5	1,7	2,4	n.a.	5,6
46	Canada	1946-2025	1,9	3,6	4,1	n.a.	2,2
47	Belgium	1946-2026	n.a.	4,2	3,1	2,6	n.a.
48	Austria	1946-2027	5,2	3,3	-3,8	n.a.	5,7
49	Australia	1946-2028	3,2	4,9	4,0	n.a.	5,9
50							
51			4,1	2,8	2,8	=AVERAGE(L30:L44)	
52							

Running Example

= D4 / D2

	A	B	C	D
1	Item	1st Qtr	2nd Qtr	Total
2	Units Sold	1000	1500	2500
3	ASP/Unit	$ 20	$ 21	$ 20,6
4	Sales Revenue	$ 20.000	$ 31.500	$ 51.500
5	Expenses	$ 5.000	$ 6.000	$ 5.000
6	Operating Income	$ 15.000	$ 25.500	$ 46.500
7	Op Income in %	75,0 %	81,0 %	90,3 %

Should be 78.6%

This is a simplified version of the homework/Budgetone spreadsheet from the EUSES Spreadsheet Corpus

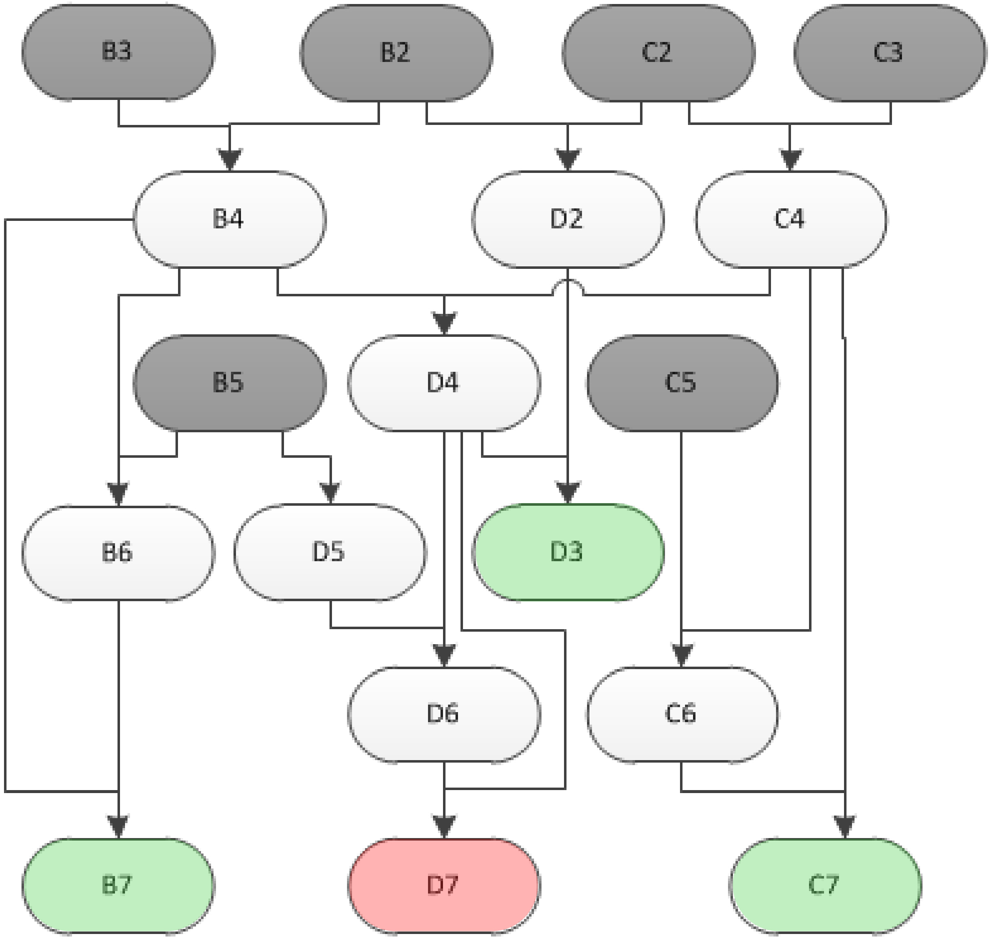
Running Example – Formula View

	A	B	C	D
1	**Item**	**1st Qtr**	**2nd Qtr**	**Total**
2	**Units Sold**	1000	1500	=SUM(B2:C2)
3	**ASP/Unit**	20	21	=D4/D2
4	**Sales Revenue**	=B3*B2	=C3*C2	=SUM(B4:C4)
5	**Expenses**	5000	6000	=SUM(B5:B5)
6	**Operating Income**	=B4-B5	=C4-C5	=D4-D5
7	**Op Income in %**	=B6/B4	=C6/C4	=D6/D4

This is a simplified version of the homework/Budgetone spreadsheet from the EUSES Spreadsheet Corpus

Running Example – Dependency Graph

	A	B	C	D
1	**Item**	**1st Qtr**	**2nd Qtr**	**Total**
2	**Units Sold**	1000	1500	=SUM(B2:C2)
3	**ASP/Unit**	20	21	=D4/D2
4	**Sales Revenue**	=B3*B2	=C3*C2	=SUM(B4:C4)
5	**Expenses**	5000	6000	=SUM(B5:B5)
6	**Operating Income**	=B4-B5	=C4-C5	=D4-D5
7	**Op Income in %**	=B6/B4	=C6/C4	=D6/D4



This is a simplified version of the homework/Budgetone spreadsheet from the EUSES Spreadsheet Corpus

Model-Based (Software) Debugging

Spreadsheet + Expect. Output → Conversion → Constraints (Model) → Solve → Diagnoses

	B	C	D
1	**1st Qtr**	**2nd Qtr**	**Total**
2	1000	1500	=SUM(B2:C2)
3	20	21	=D4/D2
4	=B3*B2	=C3*C2	=SUM(B4:C4)
5	5000	6000	=SUM(B5:B5)
6	=B4-B5	=C4-C5	=D4-D5
7	=B6/B4	=C6/C4	=D6/D4

Should be 78,6%

AB(D2) $\vee$ *behavior(D2)*
AB(D3) $\vee$ *behavior(D3)*
AB(B4) $\vee$ *behavior(B4)*
…

Test case:
D7 == 0,786
B7 == 0,750
…

For single faults:
SUM(AB(D2),AB(D3),…)==1

- Single Fault:
  - {D5}
  - {D6}
  - {D7}
- Double Fault:
  - {D3,D4}
  - …

All models are automatically derived from a faulty spreadsheet and also contain the fault!

[AHP14] Abreu, Hofer, Perez, Wotawa: "Using constraints to diagnose faulty spreadsheets." *Software Quality Journal, pp. 1–26,* 2014.

Models for a Spreadsheet's Behavior

Value-based

D2 = B2 + C2

D3 = D4 / D2

Dependency-based

$ok(B2) \land ok(C2) \rightarrow ok(D2)$

$ok(D4) \land ok(D2) \rightarrow ok(D3)$

	A	B	C	D
1	**Item**	**1st Qtr**	**2nd Qtr**	**Total**
2	**Units Sold**	1000	1500	=SUM(B2:C2)
3	**ASP/Unit**	20	21	=D4/D2
4	**Sales Revenue**	=B3*B2	=C3*C2	=SUM(B4:C4)
5	**Expenses**	5000	6000	=SUM(B5:B5)
6	**Operating Income**	=B4-B5	=C4-C5	=D4-D5
7	**Op Income in %**	=B6/B4	=C6/C4	=D6/D4

Models for a Spreadsheet's Behavior

Value-based

D2 = B2 + C2

D3 = D4 / D2

- + exact, few diagnoses
- – computation time
- – Reals: lacking support

Dependency-based

$ok(B2) \wedge ok(C2) \rightarrow ok(D2)$

$ok(D4) \wedge ok(D2) \rightarrow ok(D3)$

- + fast
- + only Boolean
- – many diagnoses

Models for a Spreadsheet's Behavior

Value-based

D2 = B2 + C2

D3 = D4 / D2

Dependency-based

$ok(B2) \wedge ok(C2) \rightarrow ok(D2)$

$ok(D4) \wedge ok(D2) \rightarrow ok(D3)$

Comparison-based

$eq(B2) \wedge eq(C2) \rightarrow eq(D2)$

$gt(B2) \wedge eq(C2) \rightarrow gt(D2)$

…

	A	B	C	D
1	**Item**	**1st Qtr**	**2nd Qtr**	**Total**
2	**Units Sold**	1000	1500	=SUM(B2:C2)
3	**ASP/Unit**	20	21	=D4/D2
4	**Sales Revenue**	=B3*B2	=C3*C2	=SUM(B4:C4)
5	**Expenses**	5000	6000	=SUM(B5:B5)
6	**Operating Income**	=B4-B5	=C4-C5	=D4-D5
7	**Op Income in %**	=B6/B4	=C6/C4	=D6/D4

Comparison-based Modeling

in1 + in2 = out

in1 * in2 = out

ABNORMAL	in1	in2	out
False	=	=	=
False	<	=	<
False	=	<	<
False	<	<	<
False	>	=	>
False	=	>	>
False	>	>	>
False	<	>	?
False	>	<	?
True	**?**	**?**	**?**

in1 - in2 = out

in1 / in2 = out

ABNORMAL	in1	in2	out
False	=	=	=
False	<	=	<
False	=	<	>
False	<	<	?
False	>	=	>
False	=	>	<
False	>	>	?
False	<	>	<
False	>	<	>
True	**?**	**?**	**?**

Practical Realization with Minion 1

```
MINION 3
# Modeling the domain = > <
# Values: 0 < / 1 = / 2 >

**VARIABLES**
DISCRETE Sheet1_F4{0..2}
DISCRETE Sheet1_F3{0..2}
DISCRETE tmp3{0..2}
DISCRETE Sheet1_H3{0..2}
DISCRETE Sheet1_B4{0..2}
DISCRETE Sheet1_D6{0..2}
DISCRETE Sheet1_B3{0..2}
DISCRETE Sheet1_D5{0..2}
DISCRETE Sheet1_D4{0..2}
DISCRETE Sheet1_D3{0..2}
DISCRETE Sheet1_F5{0..2}
DISCRETE tmp0{0..2}
DISCRETE Sheet1_B8{0..2}
DISCRETE Sheet1_B7{0..2}
DISCRETE Sheet1_B6{0..2}
DISCRETE Sheet1_B5{0..2}
DISCRETE tmp9{0..2}
BOOL ab[8]
```

```
**TUPLELIST**
plusMultFunction 40 4
0 1 1 1
0 2 1 2
0 1 2 2
0 2 2 2
0 0 1 0
0 1 0 0
0 0 0 0
0 0 2 0
0 0 2 1
0 0 2 2
0 2 0 0
0 2 0 1
0 2 0 2
1 0 0 0
1 0 0 1
1 0 0 2
1 1 0 0
1 1 0 1
1 1 0 2
1 2 0 0
1 2 0 1
1 2 0 2
1 0 1 0
1 0 1 1
1 0 1 2
1 0 2 0
1 0 2 1
1 0 2 2
1 1 1 0
```

```
minusDivFunction 40 4
0 1 1 1
0 2 1 2
0 1 2 0
0 2 2 0
0 2 2 1
0 2 2 2
0 0 1 0
0 1 0 2
0 0 0 0
0 0 0 1
0 0 0 2
0 0 2 0
0 2 0 2
1 0 0 0
1 0 0 1
1 0 0 2
1 1 0 0
1 1 0 1
1 1 0 2
1 2 0 0
1 2 0 1
1 2 0 2
1 0 1 0
1 0 1 1
1 0 1 2
1 0 2 0
1 0 2 1
1 0 2 2
1 1 1 0
```

Practical Realization with Minion 2

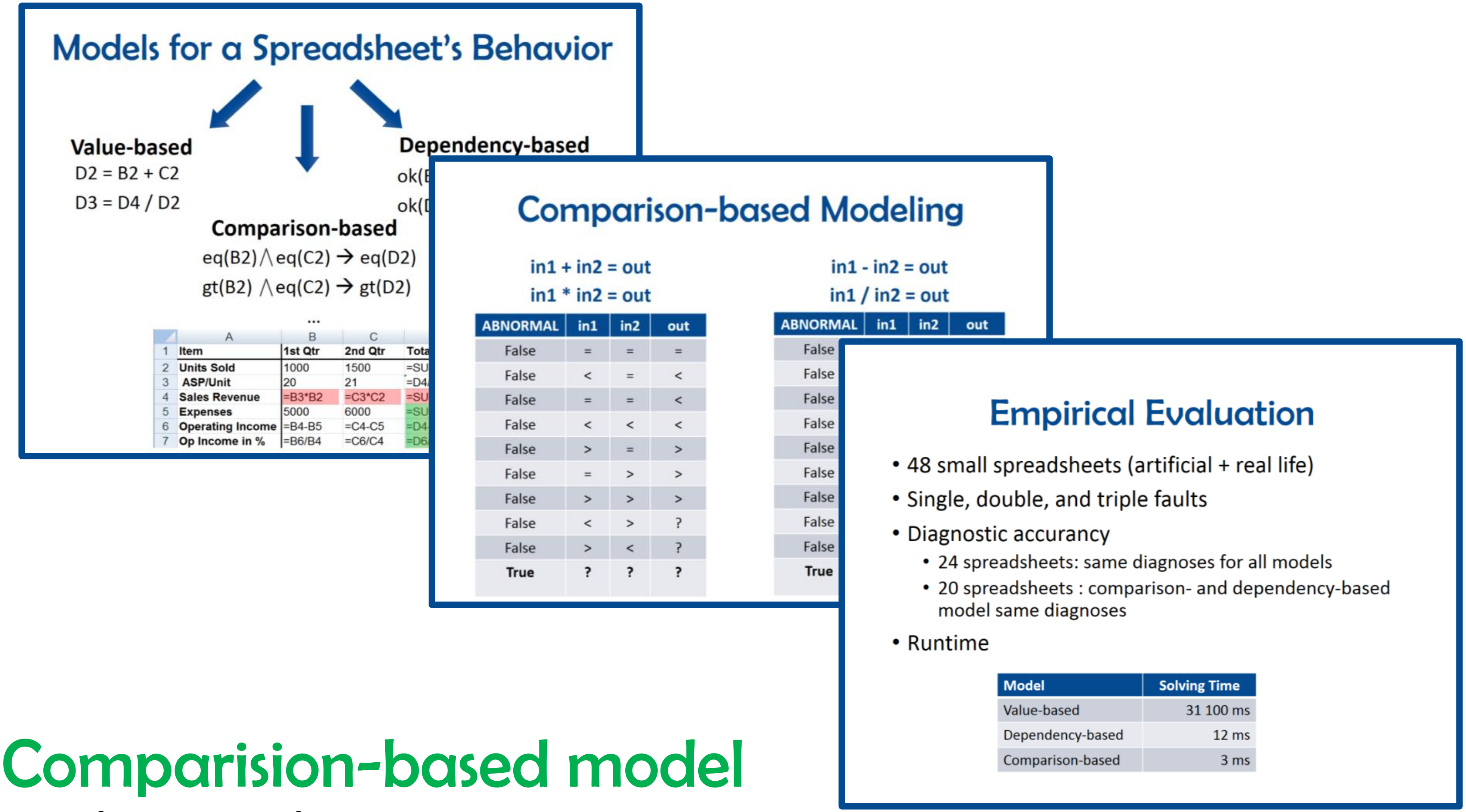
```
**SEARCH**
VARORDER [ab]
PRINT [ab]

**CONSTRAINTS**
# System description
table([ab[4],Sheet1_D4,Sheet1_D5,Sheet1_F4], plusMultFunction)
table([ab[1],Sheet1_D3,Sheet1_D4,Sheet1_F3], plusMultFunction)
table([ab[3],Sheet1_B4,Sheet1_B6,Sheet1_D4], plusMultFunction)
table([ab[5],Sheet1_B7,Sheet1_B8,Sheet1_D5], plusMultFunction)
table([ab[6],Sheet1_D5,Sheet1_D6,Sheet1_F5], plusMultFunction)
table([ab[7],Sheet1_B8,Sheet1_B6,tmp9], plusMultFunction)
table([ab[0],Sheet1_B3,Sheet1_B4,tmp0], plusMultFunction)
table([ab[7],tmp9,Sheet1_B4,Sheet1_D6], plusMultFunction)
table([ab[2],Sheet1_F3,Sheet1_F4,tmp3], minusDivFunction)
table([ab[2],tmp3,Sheet1_F5,Sheet1_H3], minusDivFunction)
table([ab[0],tmp0,Sheet1_B5,Sheet1_D3], plusMultFunction)

# TEST CASE / Observations
eq(Sheet1_B8,1)
eq(Sheet1_B7,1)
eq(Sheet1_B6,1)
eq(Sheet1_H3,0)
eq(Sheet1_B3,1)
eq(Sheet1_B4,1)
eq(Sheet1_B5,1)

#SIZE OF SOLUTION
watchsumgeq(ab,1)
watchsumleq(ab,1)
**EOF**
```

Empirical Evaluation

- 48 small spreadsheets (artificial + real life)
- Single, double, and triple faults
- Diagnostic accurancy
  - 24 spreadsheets: same diagnoses for all models
  - 20 spreadsheets : comparison- and dependency-based model same diagnoses
- Runtime

Model	Solving Time
Value-based	31 100 ms
Dependency-based	12 ms
Comparison-based	3 ms

Summary and Conclusion



Comparision-based model

- Short solving time
- Good diagnostic accurancy

→ **Useful in practice**