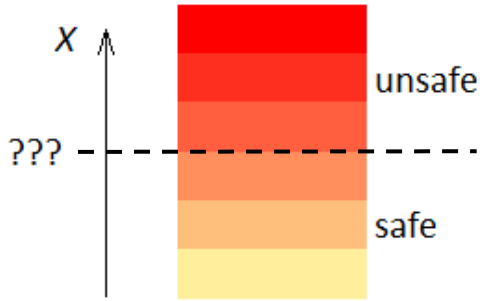
TRANSPORT RESEARCH CENTRE

Which validation is more valid?

Jiří Ambros
CDV – Transport Research Centre
Czech Republic

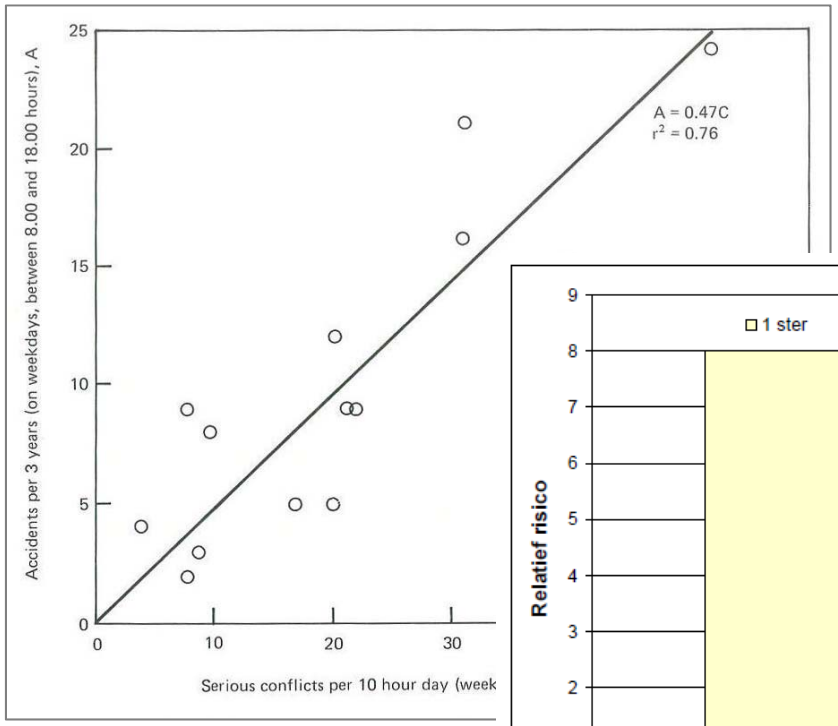
Introduction

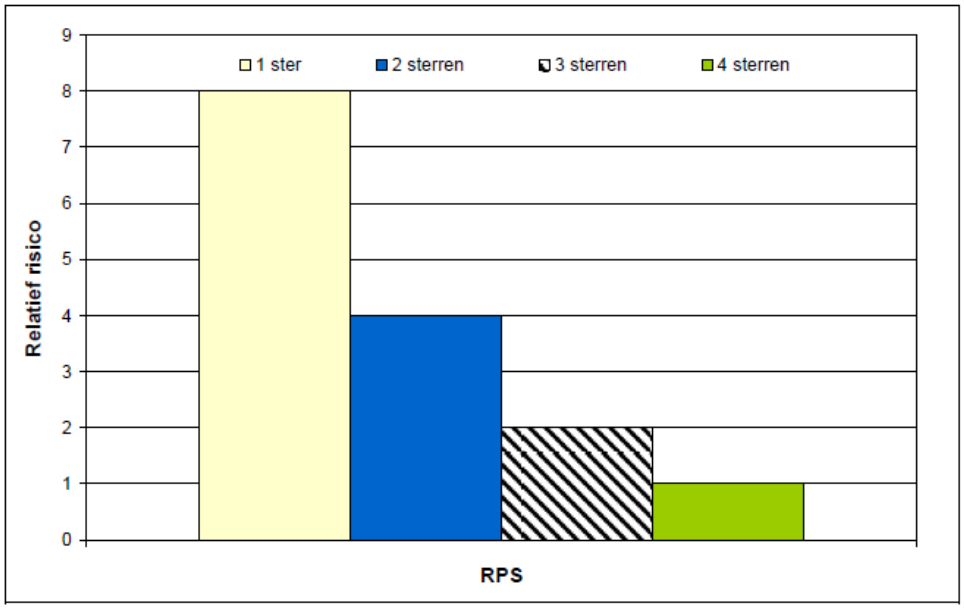


- How to collect data for validation?
- A potential surrogate *X* in road network…
  - Overall validation (correlate *X* and accident frequency)
  - Define safe/unsafe roads, collect *X*, estimate cut-off value
  - Define a cut-off value and check the relationship to safety
  - Naturalistic driving study (safe behaviour → safe *X* values)

Jiří Ambros, jiri.ambros@cdv.cz

Desired validation results...





Jiří Ambros, jiri.ambros@cdv.cz

Examples

- Floating car data
  - Company vehicle fleets, GPS + IMU
- Example analyses
  - Speeds
  - Accelerations
  - Jerks

Jiří Ambros, jiri.ambros@cdv.cz

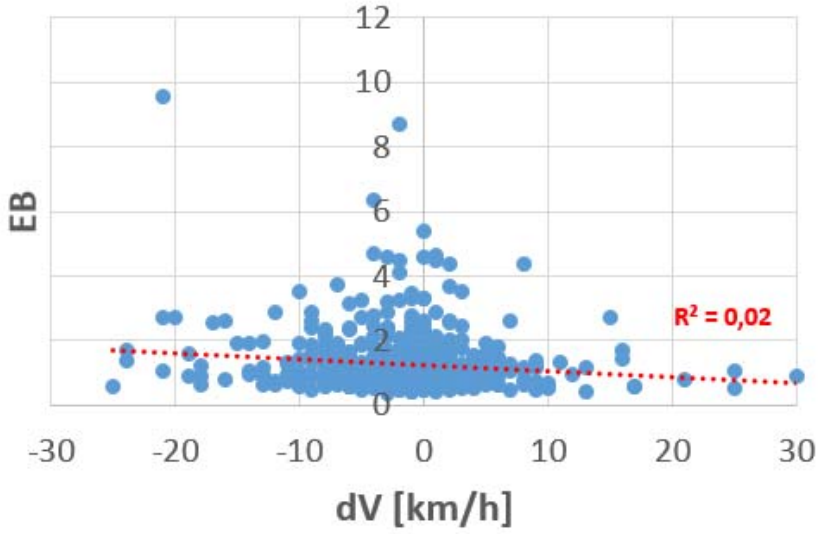
Example 1: Speed consistency

- Theory: unsafe (unexpected) curves → hard braking
  - Speed consistency $dV = V_{curve} - V_{tangent}$
  - Negative dV = braking (the less dV, the more risk)
- GPS data in 509 tangent-curve pairs (with 100+ drives)
- Safety level (6 yrs acc.) estimate adjusted by accident prediction model → empirical Bayes estimate (*EB*)

Jiří Ambros, jiri.ambros@cdv.cz

Approach 1: Overall validation

- Relationship between *dV* and *EB* ?
- No correlation

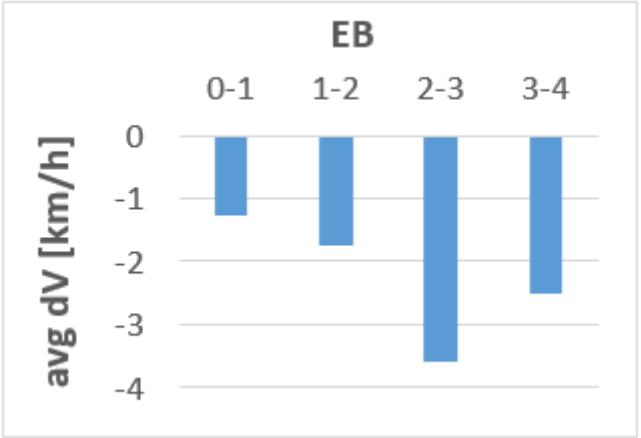


Jiří Ambros, jiri.ambros@cdv.cz

Approach 2: Safe roads → safe consistency?

- Using pivot tables
- The higher risk, the smaller sample
- Sign of trend, but no clear *dV* threshold

EB	avg dV	n
0-1	-1,27	252
1-2	-1,74	190
2-3	-3,60	42
3-4	-2,50	12
...	...	...

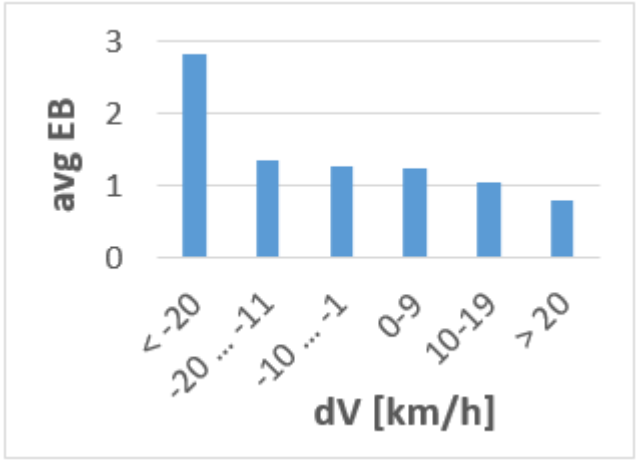


Jiří Ambros, jiri.ambros@cdv.cz

Approach 3: Is there a cut-off value of *dV*?

- Sample again limited on borders
- Cut-off at –20 km/h, consistent with past research

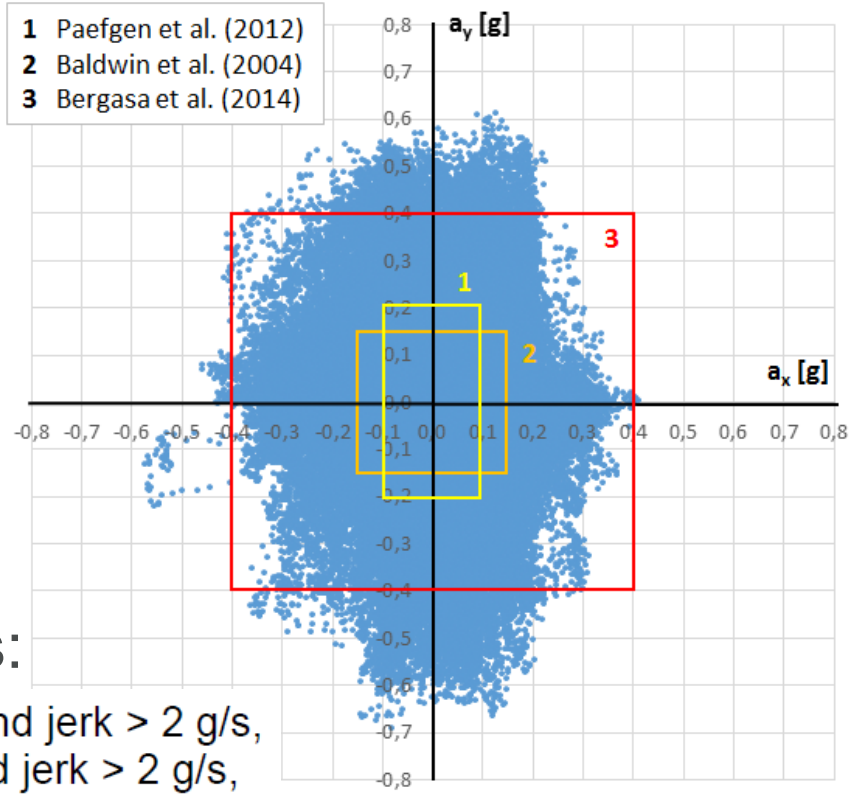
dV	avg EB	n
< -20	2,84	6
-20 ... -11	1,35	25
-10 ... -1	1,27	272
0-9	1,22	190
10-19	1,04	12
> 20	0,80	4



Jiří Ambros, jiri.ambros@cdv.cz

Example 2: Accelerations

a_x … braking/accelerating
a_y … left/right turns



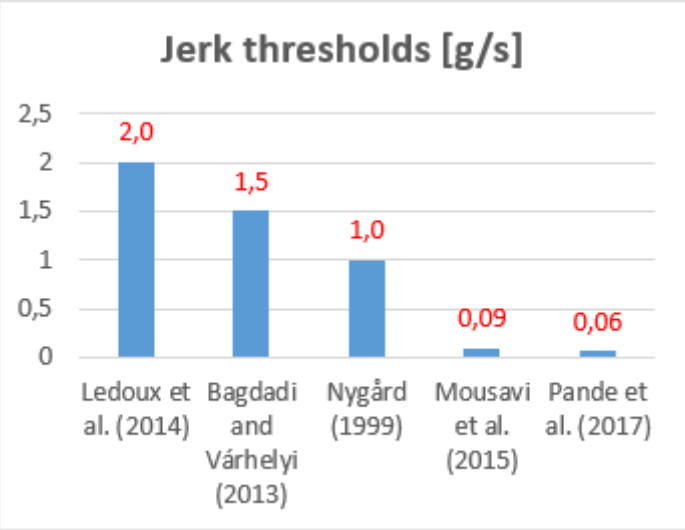
- Various risk space definitions
- Combined with speed and jerks:

- Speed <= 80 km/h and acceleration norm > 0.6 g and jerk > 2 g/s,
- Speed > 80 km/h and acceleration norm > 0.5 g and jerk > 2 g/s,
- Speed > 100 km/h and acceleration norm > 0.4 g and jerk > 2 g/s.

Jiří Ambros, jiri.ambros@cdv.cz

Example 3: Jerks

Rate of change of deceleration (*da/dt*)



- *21 jerk value thresholds were evaluated in the sensitivity analysis* (…) *The jerk-rate was then compared to the crash rate for each segment* (Mousavi et al., 2015)

- *The threshold value X was varied from 0.50 ft/s³ to 2.75 ft/s³ with an increment of .25 ft/s³* (Pande et al., 2017) … 10 thresholds

- Theory-based or data-based ?

Summary

Larger ("cheap") studies

- Using other party datasets, such as vehicle fleet data
- One can remove outliers, select subsets…

Smaller ("expensive") studies

- Not network-wide
- For example traffic conflict studies: mostly 1 site only

Jiří Ambros, jiri.ambros@cdv.cz

~~Conclusions~~ Discussion

"Which validation is more valid?"

- Validation approach depends on amount of available data
- Big data → "data mining"
- Small "expensive" data → ???
- Product / process validation

Jiří Ambros, jiri.ambros@cdv.cz

TRANSPORT RESEARCH CENTRE

Thank you for your attention

Jiří Ambros
jiri.ambros@cdv.cz

CDV – Transport Research Centre
Líšeňská 33a, 636 00 Brno, Czech Republic

www.cdv.cz/en/