Crystal Ball: Phase 2 Report (part 1) Application to NASA systems

Tim Menzies
Barry Boehm
Jairus Hihn

July 22, 2008

Roadmap

Executive Summary

Background

Baselines

Solo discussions

Group discussions

What's next?

Executive summary

- We seek clarity, consensus of the future of software at NASA
  - We seek this at a time of massive institutional change at NASA
- We did not find consensus between individuals at the detailed level
  - But we did find consensus on general trends during group discussions
- Conclusion:
  - Our software process planning tools should be aware of those trends and offer conclusions over the options space within those trends
  - E.g. the STAR/NOVA tool proposed by Menzies/Hihn/Boehm

Roadmap

Executive Summary

Background

Baselines

Solo discussions

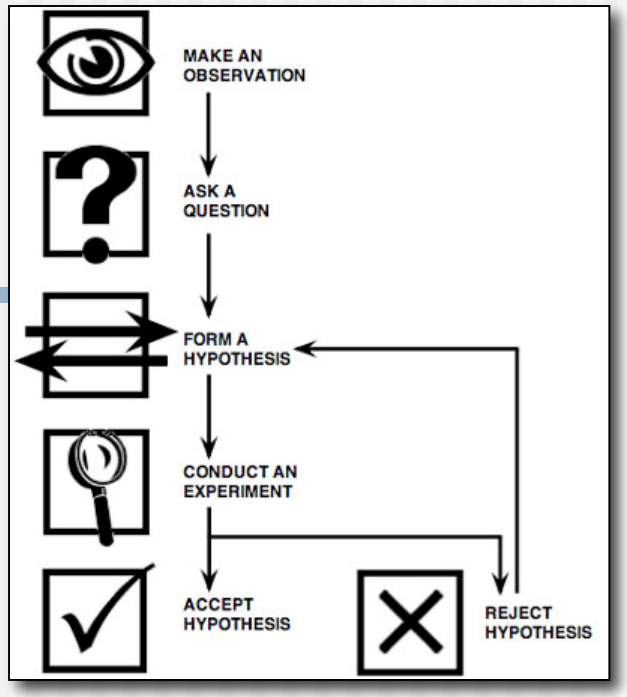
Group discussions

What's next?

Possibility

- Maybe, can use AI to better plan software projects
  - Monte Carlo, simulated annealing, Bayesian feature selection
  - Result: better peeking into the future
- E.g. Menzies, Boehm, Hihn, Lum ASE 2007
  - AI search methods on software process models
  - Some stable predictions in a huge space of options
  - http://menzies.us/pdf/07casease.pdf
- To test that possibility
  - Need ranges representing current & future NASA environment

Methodology



- 3 USC software process models
  - 1 WVU AI search engine
  - Multiple case studies
  - Sensitivity analysis
    - Found nine key factors
- Manual exploration of those factors with experienced software experts
  - May 2008: SE research leaders
  - July 2008: JPL experts
- 1 hour one-on-one sessions
  - Followed 2 days later by a 3 hour group meeting

Roadmap

Executive Summary

Background

Baselines

Solo discussions

Group discussions

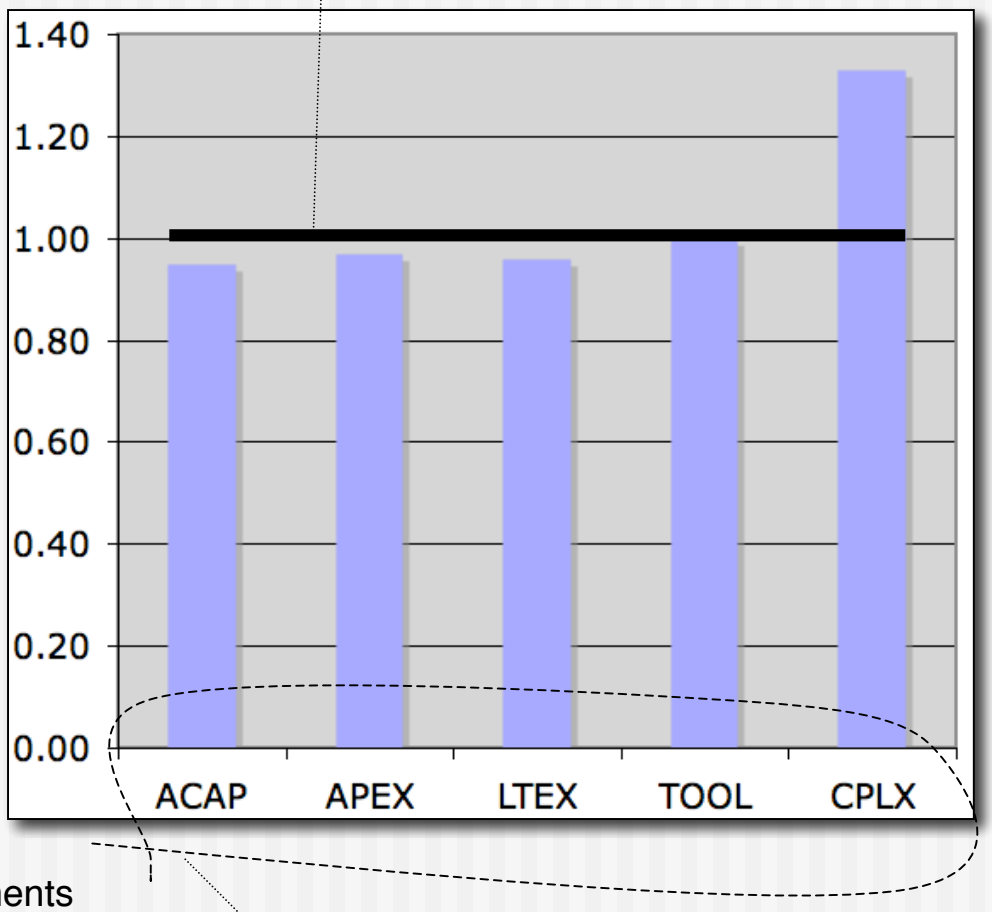
What's next?

Baselines collected from NASA data

- Some JPL
- Some from other centers
- Data collected for the decade 2000 to 2008.
  - We will refer to this data as the “2005” data

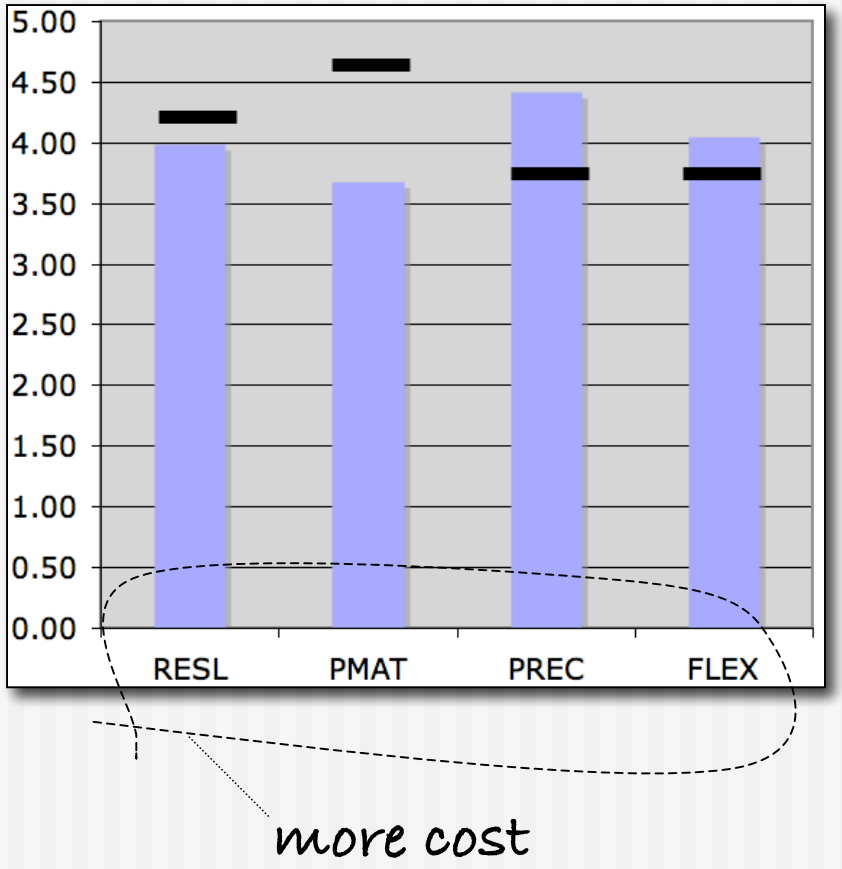
Baselines (1)

- Data from NASA systems, 2000 to 2008
- Acap (analyst capability)
  - Worse: worst 15%
  - Middle: 55%
  - Best: best 10%
- Apex (applications experience)
  - Worst: 2 months
  - Middle: 1 year
  - Best: 6 years
- Ltex (language and toolset experience)
  - Worst: 2 months
  - Middle: 1 year
  - Best: 6 years
- Tool (use of software tools)
  - Worst: edit,code,debug
  - Best: integrated with life cycle
- Cplx (product complexity)
  - Worse: e.g. simple read/write statements
  - Middle: e.g. use of simple interface widgets
  - Best: e.g. performance-critical em- bedded systems



Baselines (2)

- Data from NASA systes, 2000 to 2008
- Resl (architecture or risk resolution)
  - Worst: few interfaces defined or few risk eliminated
  - Middle: most interfaces defined or most risks eliminated
  - Best: all interfaces defined or all risks eliminated
- Pmat process maturity
  - Worst: CMM level 1
  - Middle: CMM level 3
  - Best: CMM level 5
- Prec Precedentedness
  - Worst: we have never built this kind of software before
  - Middle: somewhat new
  - Best: thoroughly familiar
- Flex (development flexibility)
  - Worst: development process rigorously defined
  - Middle; some guidelines, which can be relaxed
  - Best; only general goals defined



Roadmap

Executive Summary

Background

Baselines

Solo discussions

Group discussions

What's next?

Data collected from one-on-one sessions

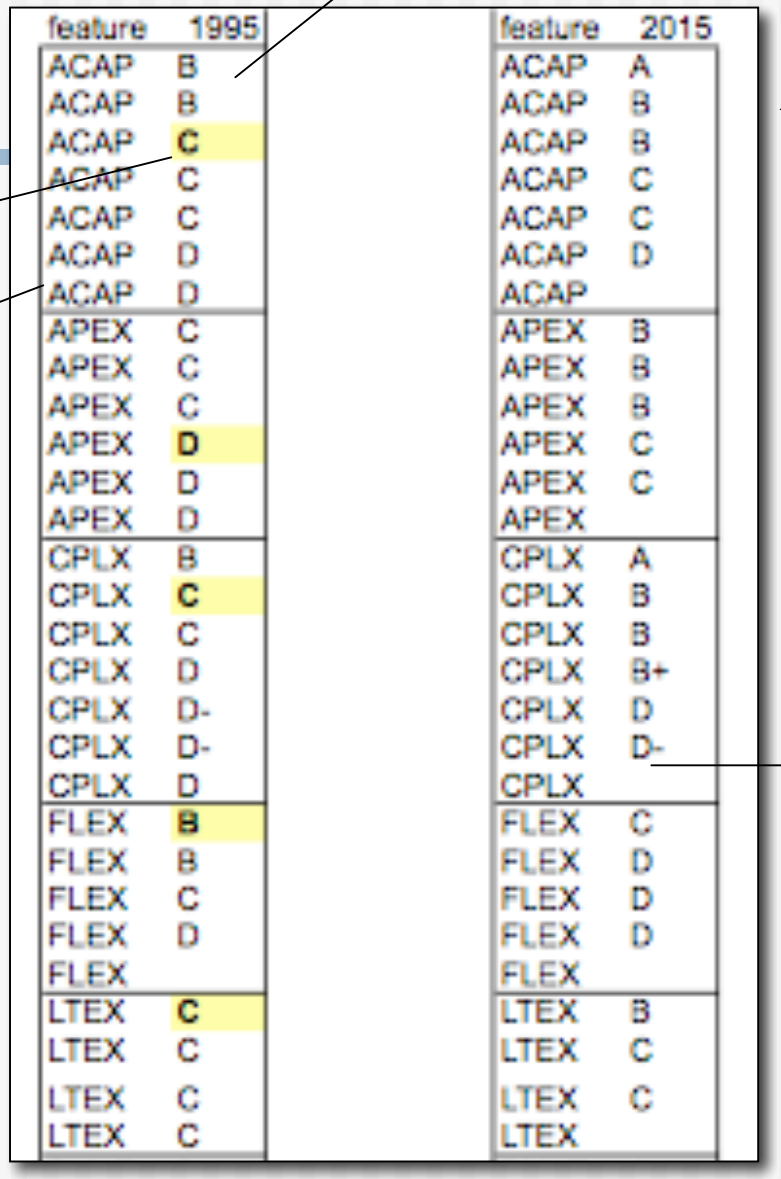
- Duration: 1 hour
- What (are we studying?)
  - JPL flight software
  - Planetary missions (e.g. rovers)
- When (are we looking)
  - -10 years (to calibrate historical data)
  - +10 years (to make predictions)
- What (are we using to express the data)
  - Projects expressed in the COCOMO ontology
- How (are we answer the questions):
  - Please answer :
    - A: significantly better
    - B: somewhat better
    - C: no change
    - D : somewhat worse
    - E : significantly worse

Results

e.g. two opinions that our analyst somewhat better In 1995 than today

Yellow shows historical record

e.g. two opinions That our analysts were worse in 1995 than today



feature	1995	feature	2015
ACAP	B	ACAP	A
ACAP	B	ACAP	B
ACAP	C	ACAP	B
ACAP	C	ACAP	C
ACAP	C	ACAP	C
ACAP	D	ACAP	D
ACAP	D	ACAP	
APEX	C	APEX	B
APEX	C	APEX	B
APEX	C	APEX	B
APEX	D	APEX	C
APEX	D	APEX	C
APEX	D	APEX	
CPLX	B	CPLX	A
CPLX	C	CPLX	B
CPLX	C	CPLX	B
CPLX	D	CPLX	B+
CPLX	D-	CPLX	D
CPLX	D-	CPLX	D-
CPLX	D	CPLX	
FLEX	B	FLEX	C
FLEX	B	FLEX	D
FLEX	C	FLEX	D
FLEX	D	FLEX	D
FLEX		FLEX	
LTEX	C	LTEX	B
LTEX	C	LTEX	C
LTEX	C	LTEX	C
LTEX	C	LTEX	

Each line is one opinion from one expert

Gaps when experts Said "don't know"

Results

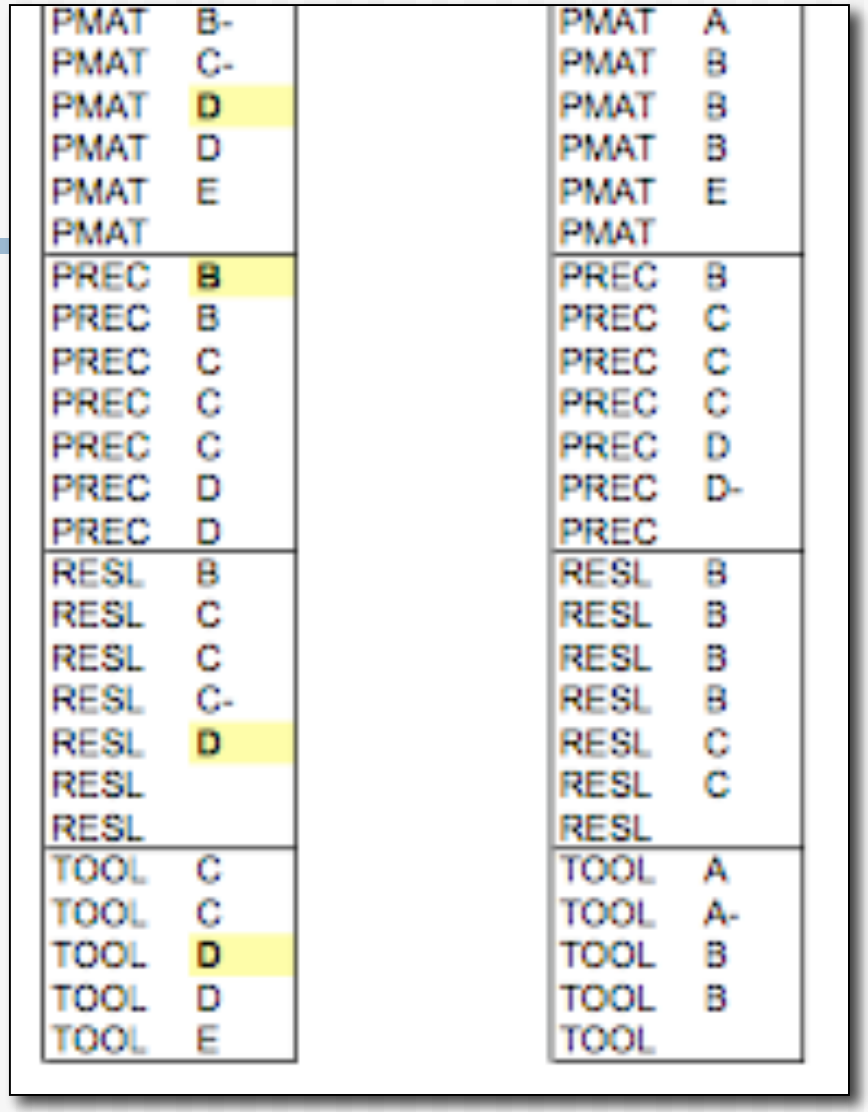
- 1995:
  - Historical record mid-point of expert view: acap,cplx
  - Historical record at one end of expert view: apex, flex
  - Historical record agrees with expert view: ltex
- 2015:
  - Usually, static or will improve
    - Acap,apex, ltex
  - No agreement on future
    - Cplx
  - Usually, static or will get worse
    - Flex
    - E.g. product lines will force uniformity

feature	1995
ACAP	B
ACAP	B
ACAP	C
ACAP	C
ACAP	C
ACAP	D
ACAP	D
APEX	C
APEX	C
APEX	C
APEX	D
APEX	D
APEX	D
CPLX	B
CPLX	C
CPLX	C
CPLX	D
CPLX	D-
CPLX	D-
CPLX	D
FLEX	B
FLEX	B
FLEX	C
FLEX	D
FLEX	
LTEX	C
LTEX	C
LTEX	C
LTEX	C

feature	2015
ACAP	A
ACAP	B
ACAP	B
ACAP	C
ACAP	C
ACAP	D
ACAP	
APEX	B
APEX	B
APEX	B
APEX	C
APEX	C
APEX	
CPLX	A
CPLX	B
CPLX	B
CPLX	B+
CPLX	D
CPLX	D-
CPLX	
FLEX	C
FLEX	D
FLEX	D
FLEX	D
FLEX	
LTEX	B
LTEX	C
LTEX	C
LTEX	

Results

- 1995:
  - Historical record mid-point of expert view: pmat, tool
  - Historical record at one end of expert view: resl, prec
    - Note: resl very hard to answer (complex question)
- 2015
  - Usually, static or will improve
    - Resl, tool
  - No agreement on future
    - Prec, pmat



PMAT	B-	PMAT	A
PMAT	C-	PMAT	B
PMAT	D	PMAT	B
PMAT	D	PMAT	B
PMAT	E	PMAT	E
PMAT		PMAT	
PREC	B	PREC	B
PREC	B	PREC	C
PREC	C	PREC	C
PREC	C	PREC	C
PREC	C	PREC	D
PREC	D	PREC	D-
PREC	D	PREC	
RESL	B	RESL	B
RESL	C	RESL	B
RESL	C	RESL	B
RESL	C-	RESL	B
RESL	D	RESL	C
RESL		RESL	C
RESL		RESL	
TOOL	C	TOOL	A
TOOL	C	TOOL	A-
TOOL	D	TOOL	B
TOOL	D	TOOL	B
TOOL	E	TOOL	

Results for reuse and size changes

- Reuse
  - Low levels in 1995
  - Usually, low levels expected for 2015
- Size of system
  - Generally perceived to be smaller before
  - Expected to grow by a factor of (around) two by 2015

	1995
reuse	0
reuse	0
reuse	0
reuse	0.02
reuse	0.1
reuse	20
size	0.2
size	0.33
size	0.5
size	0.7
size	0.8
size	D

	2015
reuse	15
reuse	0.2
reuse	0.13
reuse	0.7
reuse	0.2
reuse	0.2
size	2
size	
size	2
size	2.5
size	1.5
size	A

Discussion

- In one-on-one sessions
  - Observed variance on expectations of future software at JPL
  - (So, great need for a tool that can explore a space of options)
- Possibility
  - If we bring the experts together in one room
    - Perhaps consensus will emerge
  - Hence, our next study… (read on)

Roadmap

Executive Summary

Background

Baselines

Solo discussions

Group discussions 

What's next?

Structure of the large group meeting

- Duration: 2 hours
- For each variable
  - Have if defined by Boehm
  - Discuss it with group
  - Vote

Discussion results: acap and cplx

- Acap will be static or increase
- Cplx: will increase
  - Note: if missions get more complex while acap remains static then some of the software complexity increase will arise from sub-optimal design decisions made by analysts struggling with harder and harder problems.

Discussion results: pmat and tool

- Pmat will improve
  - But much concern over “superficial” pmat improvement (just checking the boxes, not really improving anything)
- Tools will improve
  - But much discussion on how our current tools are not fully utilized (to say the least) by the average programmer

Other discussion results

- Group discussion yielded much more consensus view that solo discussions (see below)
- But
  - Group discussions time consuming
  - We could only cover half the variables in the available time.

Roadmap

Executive Summary

Background

Baselines

Solo discussions

Group discussions

What's next?

Next steps

- While no detailed agreement on specific changes
  - No dramatic changes to historical trends
  - Some agreement on general trends
- We can now try simulating over those trends with our AI tool
  - Results to follow