

Predictor

a Prediction Support Tool

Henrik Bergström

Stockholm University/Royal Institute of Technology

Department of Computer and Systems Sciences

Electrum 230, S-164 40 Kista, Sweden

e-mail: henrikbe@dsv.su.se

Outline

- **Cost estimation in general**
- Cost estimation in Predictor
- CBR control parameters
- Concluding remarks
 - How
 - What
 - Where

Outline

- Cost estimation in general
- **Cost estimation in Predictor**
- CBR control parameters
- Concluding remarks
 - How
 - What
 - Where

Outline

- Cost estimation in general
- Cost estimation in Predictor
- **CBR control parameters**
- Concluding remarks
 - How
 - What
 - Where

Outline

- Cost estimation in general
- Cost estimation in Predictor
- CBR control parameters
- **Concluding remarks**
 - **How**
 - **What**
 - **Where**

Background

- Cost Estimation of Software Projects
 - Necessary
 - Hard

Prediction Assistance for Project Management

- 1998-2000
- Swedish National Board for Industrial and Technical Development – NUTEK
- Project Software Sweden AB
- IBM Sweden AB
- Ericsson Telecom AB

Goals of the project

Develop prediction models for project planning and monitoring based on local measurement data, and to implement them in a project management tool set.

Algorithmic Models

- General Form

$$\text{Effort} = \alpha \cdot S^{\beta} + c$$

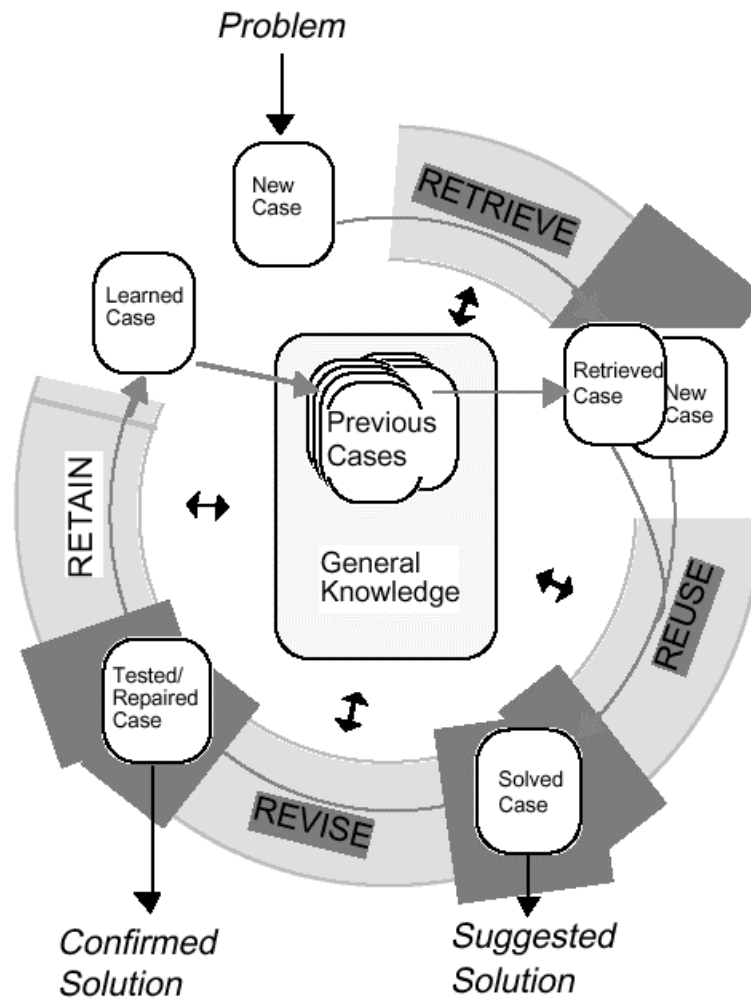
S : the size of the project

α
 β
 c } Constants

Performance of Algorithmic Models according to [Kemerer87]

- Slim $\approx 771\%$ estimation error
- COCOMO $\approx 600\%$ estimation error
- Function Points $\approx 100\%$ estimation error
- Estimacs $\approx 85\%$ estimation error

Case-Based Reasoning



Aamodt, A. and Plaza, E., "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *AI Communications*, vol. 7(1), pp. 39-59, 1994.

Predictor - Why

- Methods developed using local data should be better than methods developed from some other organisations data

Predictor - Functionality

- Measurement database
- Estimation techniques
 - Expert judgement
 - Local algorithmic models
 - Case-based reasoning
- Experience database

Predictor - Intended Users

- Software project managers
- Other
 - Managers in general
 - Developers
 - Customers
 - Process improvement

Predictor - User Interface

The screenshot shows the Predictor v 1.0 - exampledb application window. The interface is divided into several sections:

- Current Project:** This section contains input fields for project details. The first entry (ID: K 1) has the following values: Lang: Cobol, Machine: IBM308X, Months: 17, MM: 287, KSLOC: 253.6, FP: 1217.1, UA FP: 1010, and LOC per M: 884. The second entry (ID: K 2) has: Lang: Cobol, Machine: IBM43XX, Months: 7, MM: 82.5, KSLOC: 40.5, FP: 507.3, UA FP: 457, and LOC per M: 491.
- Prediction:** This section has tabs for CBR Prediction, Analogies, and Algorithmic. Under CBR Prediction, there are sub-tabs for Information and Graphic. The Information tab is active, displaying the text: "Welcome to Predictor - Your prediction support tool".
- Earlier Predictions:** This section displays a table of previous predictions.

Annotations on the left side of the image:

- A bracket labeled "The project we are currently working on" points to the first entry in the Current Project list.
- A bracket labeled "Project database" points to the entire Current Project list.

Annotations on the right side of the image:

- A bracket labeled "Information about the current prediction" points to the Prediction section.
- A bracket labeled "Information on earlier predictions" points to the Earlier Predictions table.

Date	Attribute	Prediction
1999-05-25	MM	287

Predictor - User Interface

elp

▶ Σ [Table Icon] [Line Graph Icon] [Bar Graph Icon] [Help Icon] [Icon]

Prediction

Information | Graphic

CBR Prediction | Analogies | Algorithmic

Project: Project 4
Attribute: Dev Time
Method: Geometrical mean
Threshold: 50
Size Adj: Yes (Size)
Prediction: 39.5

Save

4

Attribute: Dev Time
Method: Geometrical mean
Threshold: 50
Size Adj: Yes (Size)
Made: 18
Total: 20
ME: -1.54
MAE: 13.21
MRE: 0.23

Predictor - Algorithmic Models

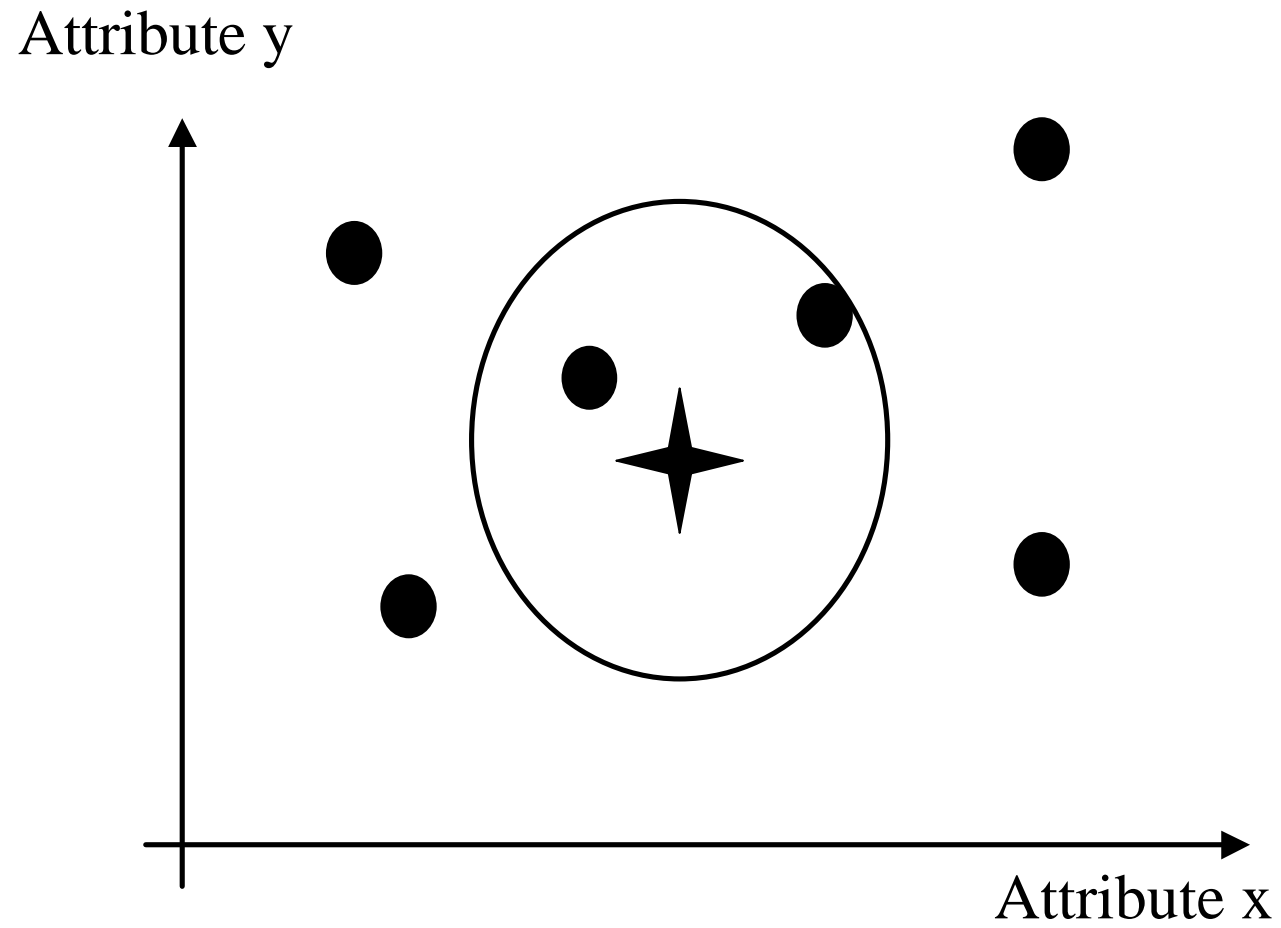
$$\text{Effort} = \alpha \cdot S^{\beta} + c$$

$$\text{Effort} = \alpha_0 + \alpha_1 V_1^{\beta_1} + \dots + \alpha_n V_n^{\beta_n}$$

Predictor - CBR

$$WED = \sqrt{\sum_{i=1}^n \left(D(p_i, q_i) \cdot W_i \right)^2}$$

Predictor - CBR



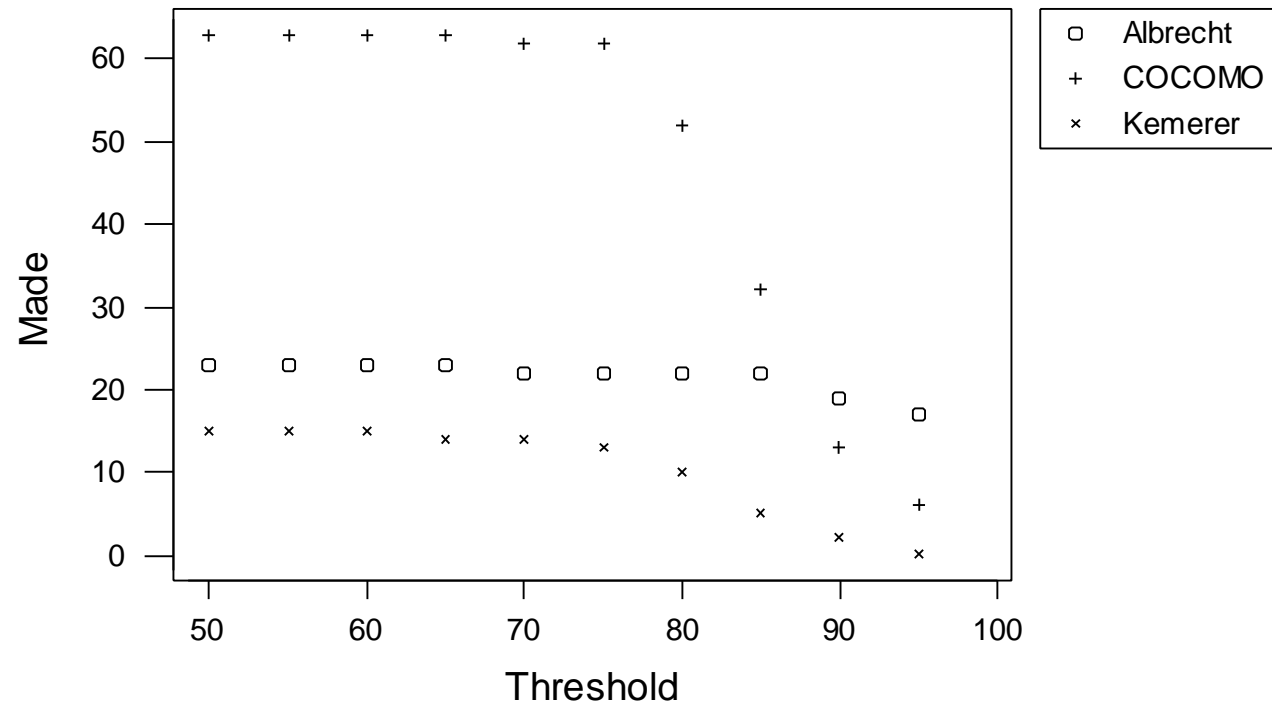
Estimation Accuracy

- Mean Relative Error (MRE)
 - The mean of the estimation error normalized for size. Should be as small as possible.
- Pred(25)
 - The percentage of the estimations that fall within 25% of the actual value. Should be as large as possible.

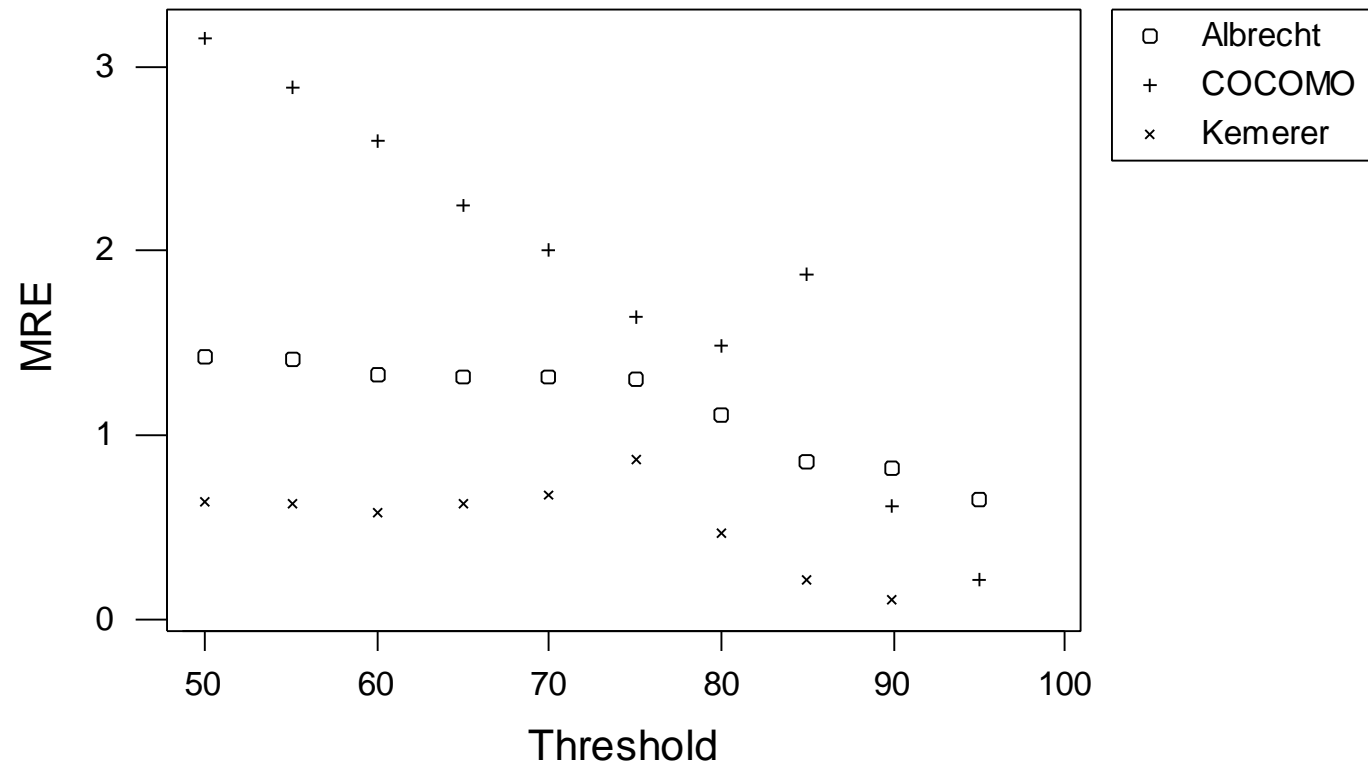
CBR Control Parameters

- At what level should the similarity threshold be set?
- Is there an adaptation function that we could use to improve the CBR predictions?

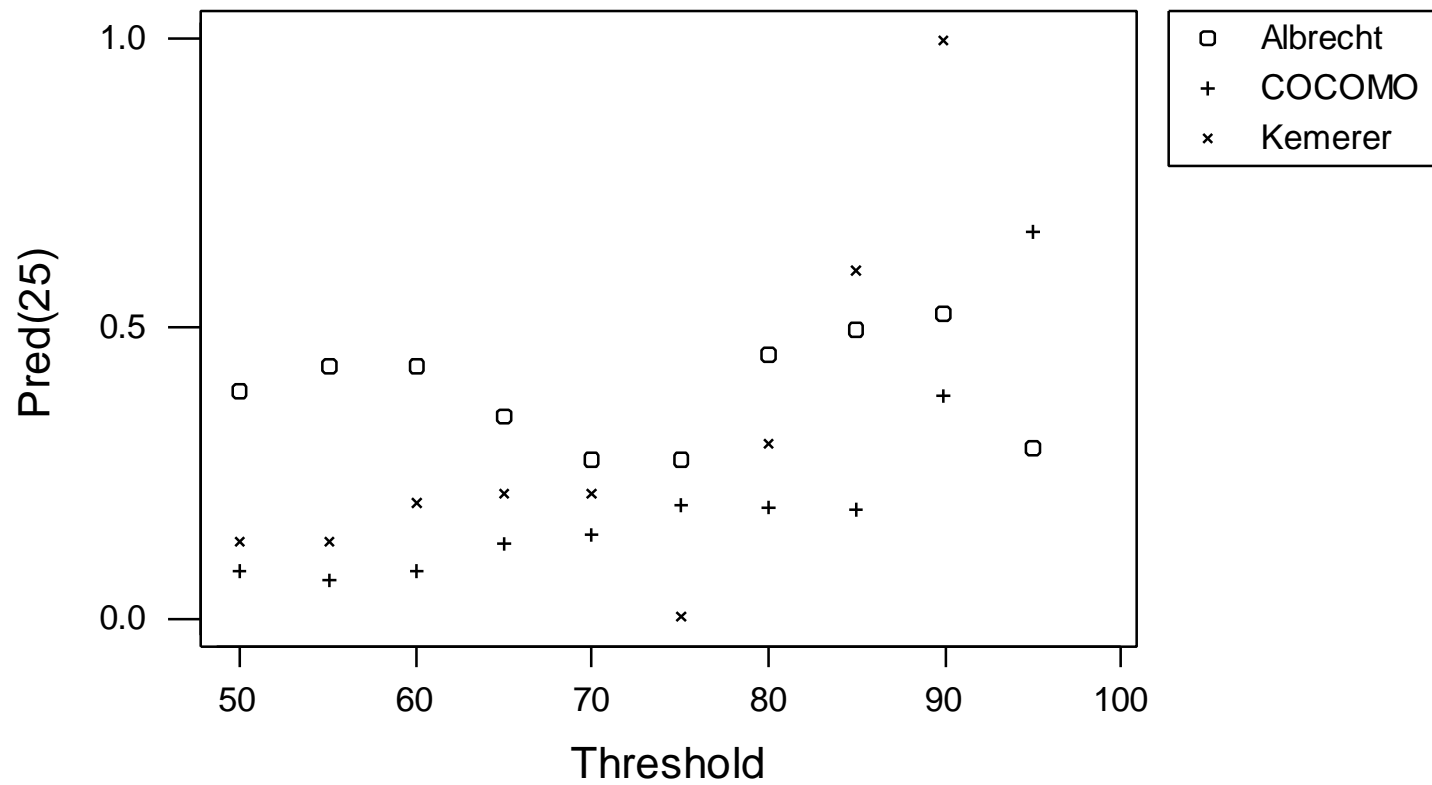
Similarity threshold vs. the number of predictions made



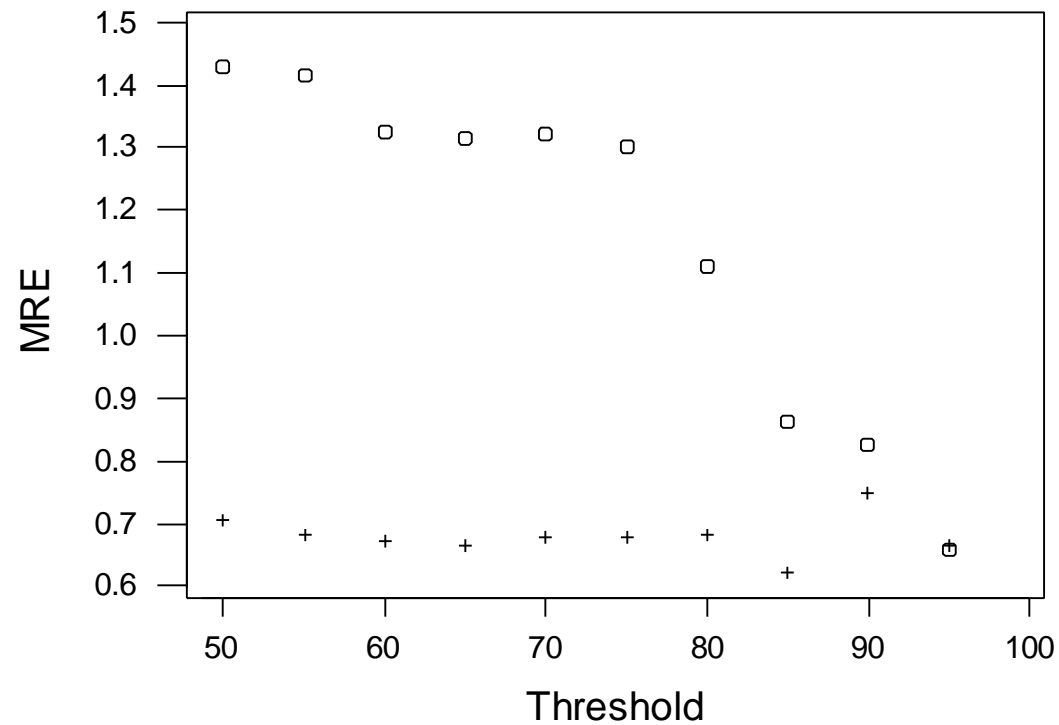
Similarity threshold vs. MRE



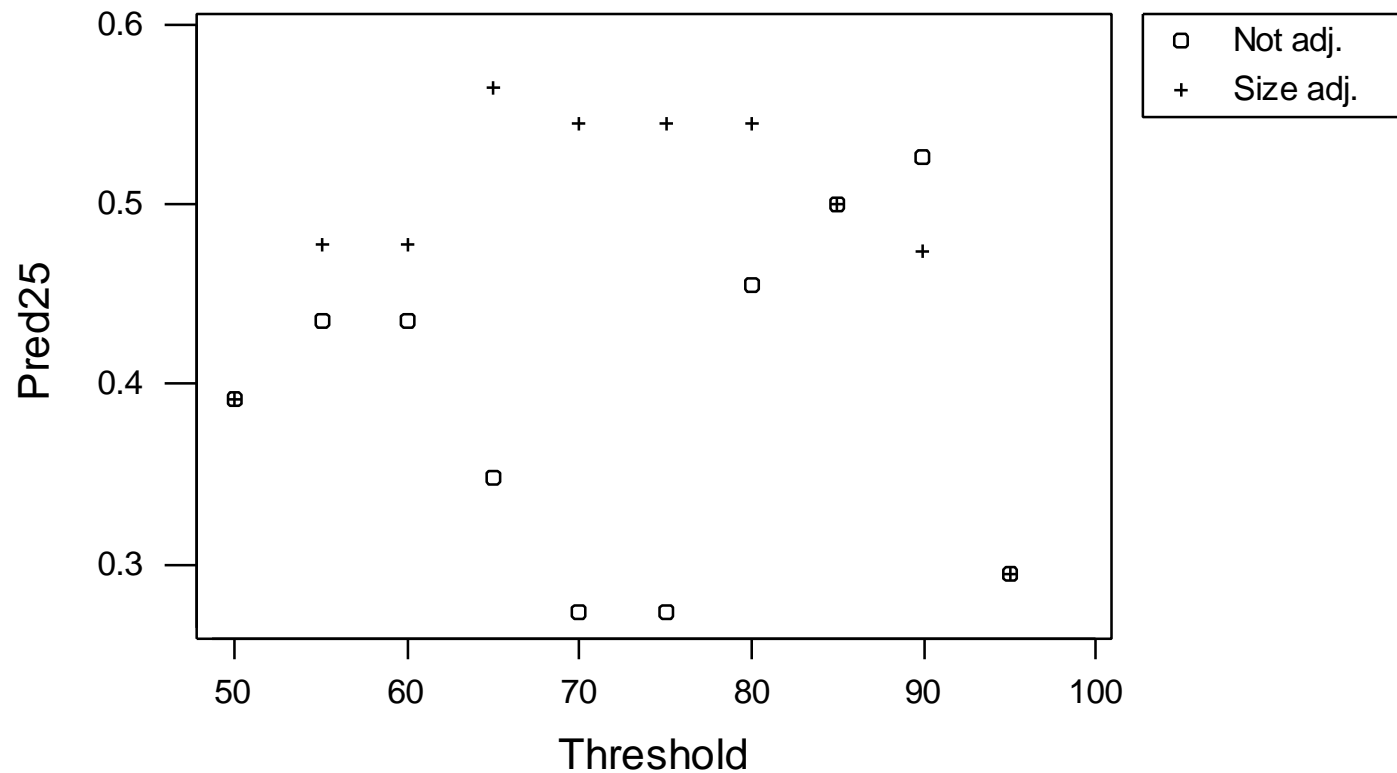
Similarity threshold vs. Pred(25)



Threshold vs. MRE for the Albrecht dataset



Threshold vs. Pred(25) for the Albrecht dataset



Predictor - Current Status

- v1.0 available
- 250+ downloads
- ??? users

Future Work

- Studying specific aspects of it
 - Attributes
 - Settings
 - Functionality

More Information

- Predictor - a Prediction Support Tool
- <http://www.dsv.su.se/~henrikbe/Predictor>
 - Program
 - Manual