Chapter 8
Time Predictions: Matching the Method to the Situation

There are many time prediction methods and principles. How should we choose between them? Time prediction methods have advantages and disadvantages that depend on the situation, but there is a scarcity of useful guidelines on how to select time prediction methods. We attempt to provide some guidance in Fig. 8.1, with more detailed explanations of the questions and the time prediction method selection advices in the subsequent paragraphs.

1. **Do you have access to historical information about one or more similar previously completed tasks?**

If you have access to data on time usage for similar tasks, you may use analogy-based time prediction methods (also known as *reference class forecasting*): Compare the characteristics of the new task with previously completed similar tasks and use the time spent on similar tasks as the predicted time on the new task. Studies on the use of analogy-based time predictions suggest that this approach mainly leads to high accuracy when the task to be predicted and one or more of the previously completed tasks are *very* similar [1]. If no data on very similar tasks are available but data on fairly similar tasks exist, the median or trimmed mean value of a larger set of fairly similar analogies (or the mean when only few analogies are identified) usually provides a good prediction.

2. **Do you have information about previously completed tasks and evidence supporting a stable relation between task characteristics and time usage?**

Information about past tasks may be used to create simple formal models of the relations between the task's characteristics and time usage. Such models may work well when there are good measures of the task's key characteristics (such as its size and complexity), when these characteristics are known before the task starts, and when we have data about key performance variables, such as the productivity of those who will complete the task. In addition, it is essential that the relations included in the model, such as that between task size and time spent, be reasonably stable. This does not mean that variations in productivity over time have to be problematic. For instance, when the average time to replace a window (say, 40 minutes) is stable, it

T. Halkjelsvik and M. Jørgensen, *Time Predictions*, Simula SpringerBriefs on Computing 5, https://doi.org/10.1007/978-3-319-74953-2_8

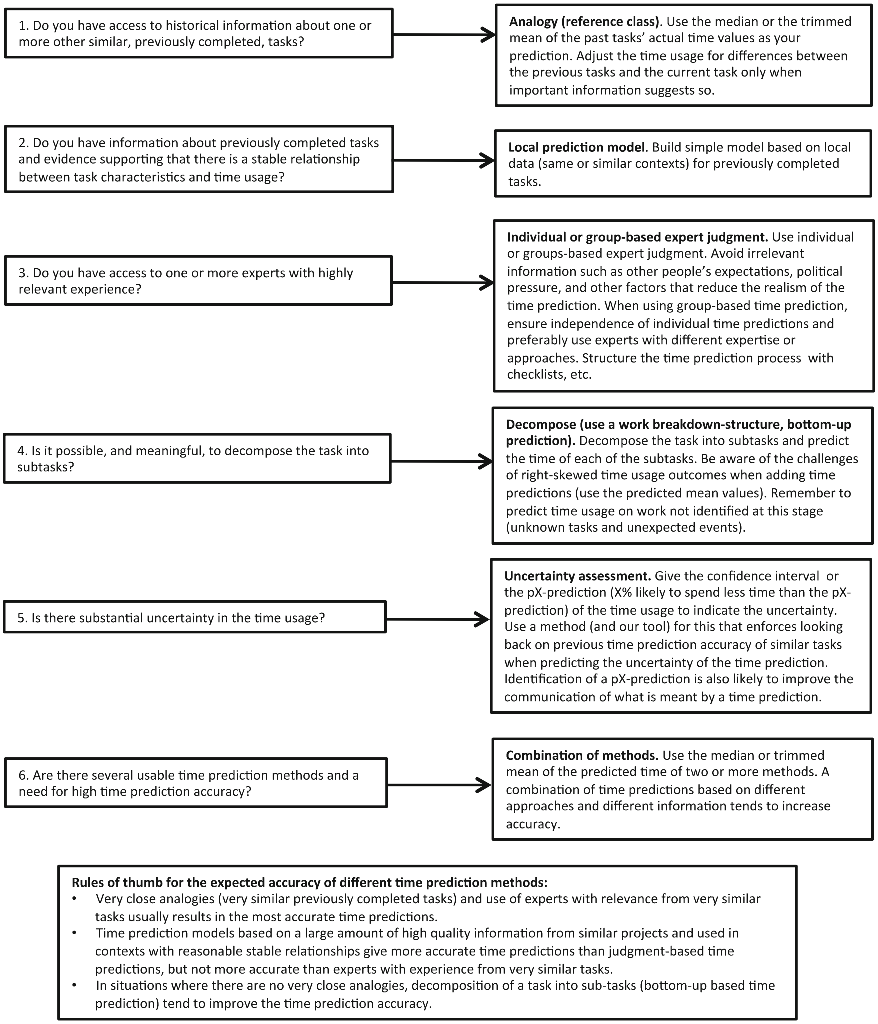

Fig. 8.1 A guide to selecting time prediction methods

does not matter whether it sometimes takes 20 minutes and at other times 90 minutes if the job is to replace 500 windows. Across many occasions, noise (e.g. bad or good luck) cancels itself out and your future average time usage will probably be close to your past average time usage, given a similar type of work and similarly skilled workers.

We are not aware of any general formal—often termed parametric—model that is valid across many time prediction contexts. The likely reason for this is the lack of highly predictive variables and stable time usage–related relations across different situations. It is possible but unlikely that we will come across examples of successful general formal time prediction models in the future and you should currently be highly suspicious if someone claims that their time production model is useful across a large number of situations. Models derived from past time usage data in your *local* context work better. Simple models are also often more accurate than complex models, probably due to less overfitting to the historical data.

3. **Do you have access to one or more experts with relevant experience?**

Expert judgements by people with *relevant* experience seem, on average, to be at least as accurate as predictions derived from formal time prediction models [2]. Be aware that the length of experience is a poor indicator of time prediction accuracy and that it is the *relevance* of the experience that matters most. This relevance quickly decreases with lack of similarity between past work and the task to be predicted.

When using expert judgement–based time predictions, it can be useful to accomplish the following:

(a) Remove irrelevant and misleading information, including information that may lead to social/organizational/political pressure, *before* sending the task information to those responsible for predicting time. Attempts to debias people exposed to irrelevant or misleading information, for example, by presenting more realistic time prediction anchors to counter the misleading one, may help but are usually not sufficient to remove the effect of the misleading information [3].
(b) Provide support and structure by using documented time prediction processes (e.g. planning poker), checklists [4], and templates for work breakdown structure.[1]
(c) Require time predictions to be justified by references to previous time usage.

4. **Is it possible and meaningful to decompose tasks into subtasks?**

Time predictions based on decomposition may give the most accurate time predictions and be the only viable option in contexts where very similar analogies cannot be found [1]. To enable accurate decomposed time predictions, there are several issues to be aware of, as follows:

- It is easy to forget to predict time spent on subtasks or activities that you are unaware of when you predict time, that is, on *unknown subtasks and unexpected events*. The best prediction of this is the proportion of time spent on such activities in the past.
- The right-skewed nature of time usage means that *adding* the time predictions of subtasks to find the total time usage should *not* be based on predictions of the *most likely* time usage but, instead, on the predicted *mean* time usage for each subtask. This requires information about the uncertainty of time usage.
- Some subtasks may be better predicted as a proportion of other subtasks rather than directly. For example, the time required for administration may be predicted as a proportion of the non-administrative tasks. Remember that the proportion of administration increases with increasing team and project size.

[1]For a general guide, see [5].

- Risk analysis, that is, the identification and assessment of events that may affect the time usage, should be an integral part of the time prediction process, not a separate activity. The risk analysis should be used as input to calculate the expected amount of extra time usage due to the risk factors.
- Dependencies between activities and between risks should be modelled. This includes modelling multiplicative dependencies between activities and correlations between risks. Monte Carlo simulations may be used for this purpose.[2]

5. **Is there substantial uncertainty in the prediction of time usage?**

Include an assessment of the uncertainty of the time usage, for instance, by the use of time prediction intervals. This will make it easier to communicate the range of possible time usage outcomes and is useful for planning and budgeting purposes. Uncertainty assessments may also be used to give predictions a more precise meaning through pX time predictions (e.g. p80 equals an 80% chance of not exceeding the prediction). It is essential to avoid uncertainty assessment methods that lead to strongly overconfident (too narrow) time prediction intervals, since they may do more harm than good.

6. **Are there several feasible time prediction methods and a need for high time prediction accuracy?**

The accuracy of time predictions may be improved by combining time predictions from different sources and using different prediction methods. *Mechanical* combinations using the median, the trimmed mean (excluding a proportion of the lowest and highest predictions), or the mean (when few time predictions are available and there are no outliers) of time predictions from different sources are possible methods for this. A proper *group-based* combination is typically based on a process in which experts individually predict the required time and then discuss and share these predictions and their rationale. Based on the discussion, the experts may choose to update their time predictions. In the final stage, the group agrees on a time prediction or uses a mechanical combination of the individual predictions. Especially for tasks where the discussion and sharing of knowledge are essential for accurate predictions, group-based combinations are likely to improve time prediction accuracy more than mechanical combinations are. The benefit of combining time predictions, regardless of the combination approach, strongly depends on the independence of the individual time predictions. If there are good reasons to weight one method more than the other, one may choose to do so; otherwise, it is better to weight all the time predictions equally.

[2]This can be done in the free software Riscue: http://www.riscue.org/.

References

1. Jørgensen M (2004) Top-down and bottom-up expert estimation of software development effort. Inf Softw Technol 46(1):3–16
2. Jørgensen M (2007) Forecasting of software development work effort: evidence on expert judgement and formal models. Int J Forecast 23(3):449–462
3. Løhre E, Jørgensen M (2016) Numerical anchors and their strong effects on software development effort estimates. J Syst Softw 116:49–56
4. Jørgensen M, Moløkken K (2003) A preliminary checklist for software cost management. In: Proceedings of IEEE third international conference on quality software, 2003. pp 134–140
5. NASA (2016) Work Breakdown Structure (WBS) Handbook. ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160014629.pdf