Chapter 5 Performance Predictions for Scheduling (14 Jan)

Again you need clear statement of purpose of the chapter in the introduction before you get to the sections component. My suggestion is that you move the first para of the methodology section to be the first para of this chapter. This has "In this chapter..". Stop at the end of this sentence and start a new paragraph.

In the new paragraph step back and talk about why you might want to make accurate predictions (some of what you currently have in the very first chapter).

Then have a third para that talks about the features and how you propose to use them. You need to remind the reader what the features are (presumably the points on the kiviat diagrams). Then a sentence or two that walks through exactly what you envisage this to work. You have much of this in what follows the "in this chapter sentence".

You say "A major benefit…developer need only instrument a kernel once…embedded as a comment". Two comments. I don't know what "instrument a kernel" means – largely because I don't think you covered this in chapter 4. What I sent you as followup email – exactly what do I need to do to use AIWC and what are the overheads. Second, I don't know what it means to have AIWC metrics embedded as a comment. This is probably something you leave to later in the chapter, possibly in the discussion/conclusion.

This para goes on to say "low overhead" and "computationally inexpensive" – (see above) neither of which you have provided any data to demonstrate!

In the para associated with the above you let slip that you are going to use the "random forest model". This needs to be elevated to a discussion in a separate paragraph. You have two things to discuss. Your requirements/goals/objectives, then options for how you might meet those requirements. So having outlined what you want to achieve, NEW PARA, what options did you consider, why are you going forward with this particular option.

You need to better outline the major sections – what can I expect to read about.

Section 5.1

The methodology heading should now be just before section 5.1.1. But I don't think "Methodology" is the right title. May be "model development".

Before section 5.1.1 you now need a para talking about what is in section 5.1

Section 5.1.1
You CANNOT say "following the same methodology outlined in ref 127" where ref 127 is YOUR paper. The methodology should be reported here. (This chapter would benefit from this being included – you need to show that you have done a lot of work!)

You need to state how many kernel results there are in total. That is something like, 4 problem sizes * Number of devices * Number of dwarfs * average kernels/dwarf. (You have this number – detail it – I call it XXX here)

You say "…in Table 3.1" – then have several sentences regarding cache sizes etc. This table is NOT in this chapter. You can reference it, but there should NOT be extensive following discussion. That is a footnote to the table if necessary, or is something that has already been discussed in chapter 3.

Section 5.1.2
Currently problematic in that there is no clear statement of what the performance model is that you are trying to build (why you need to raise the profile of why you chose RFM. Suggest you probably call this section something like "constructing the RF performance model"

Fig 5.1 – state what predicted error is for. Presumably execution time for all XXX benchmarks.

Section 5.1.3 – tittle again – maybe "Parameters for the RF performance model"

Section 5.1.4 – should this be part of the evaluation? The title suggests it is more evaluation. I would have more general title like "Training the RF performance model"

Section 5.2

Fig 5.4 – Useful to know how many data points (i.e. what is XXX). Do you include the 45degree line, maybe even lines that indicate less than 5% error?

Section 5.2.1
Last para you comment on init_alpha_dev. Is there anything special about the instruction mix in this kernel, are the problems linked to particular hardware?

Section 5.2.2
Ideally I think para 2 should be before para 1, but you would need to reword. Para 2 is about purpose. What are you trying to do. Given a purpose or objective, then comes the experiment to assess that purpose.

Current para 1 says "4 randomly selected kernels". You should be very careful with use of "random". It suggests a lack of purpose or thought. Much better is to select for specific reasons.

The comments about only needing to predict relative order correctly are good.

Fig 5.5 – why do the tows with the dark black boxes not line up between the Tiny and Small? Why are there fewer kernels for the medium and large. (I think I commented on this elsewhere in your thesis. You need to remind the reader why.)

Section 5.2.3 (NOT CURRENT)
You need a section that talks about the cost of making a prediction. Also it would be helpful for you to walk through how random person would use this. For example – Giuseppe is working on the GAMESS quantum chemistry package. This is a very large packaged (millions of lines). It includes parts that run on GPUs which – for the point of argument – you can assume are written in OpenCL. The OpenCL kernels are not one of your dwarfs. How do I use your software with GAMESS in order to direct OpenCL kernels to the optimal device. (Note that I have told you nothing about the devices).

Section 5.3
Discussion section is too brief. Include more conclusions and discussion. Critic what you have done.