

An Indian-Australian research partnership

<b>Project Title:</b>	<b>Data Mining to Improve Water Management</b>
<b>Project Number</b>	IMURA0261
Monash Supervisor(s)	Dr. Terence Chan, Prof. Shonali Krishnaswamy, <i>Full names and titles</i>
Monash Primary Contact:	<a href="mailto:terence.chan@monash.edu">terence.chan@monash.edu</a> , +61 3 9905 4079 <a href="mailto:Shonali.Krishnaswamy@monash.edu">Shonali.Krishnaswamy@monash.edu</a> , +61399031967 <i>Email, phone</i>
IITB Supervisor(s)	Prof. Ganesh Ramakrishnan, <i>Full names and titles</i>
IITB Primary Contact:	<a href="mailto:ganesh@cse.iitb.ac.in">ganesh@cse.iitb.ac.in</a> , +91 22 2576 7728 <i>Email, phone</i>

## Research Academy Themes:

Highlight which of the Academy's Theme(s) this project will address?

(Feel free to nominate more than one. For more information, see [www.iitbmonash.org](http://www.iitbmonash.org))

1. **Advanced computational engineering, simulation and manufacture**
2. Infrastructure Engineering
3. Clean Energy
4. **Water**
5. Nanotechnology
6. Biotechnology and Stem Cell Research

## The research problem

*Define the problem*

With respect to the Computational Engineering Theme, the problem is that Data mining in complex systems is difficult. There are spurious relationships, co-correlations, incomplete datasets, etc. Validation of the outcomes of data mining is difficult and often does not occur – there is need for an expert input validation phase.

With respect to the Water Theme, the problem is that natural systems such as catchments/watersheds are highly complex and difficult to manage. Recently, graphical computer models such as Bayesian networks, based on expert conceptualisation of the systems have been one method we use to attempt to simplify the system and enhance our understanding and management of it. However, the complexity of these systems also means that despite large datasets, our current knowledge and understanding of these systems is incomplete. In particular, there is likely to be bias/prejudice/dominant-paradigm-acceptance in expert opinion of environmental relationships/processes as elicited for conceptual diagrams/BN structures, but this bias is difficult to identify.

## Project aims

*Define the aims of the project*

The problems described in the section above have had some research and partial solutions applied within each discipline. The aim of this project is to combine these complementary approaches, comparing top-down (expert elicited) model structures from the field of water management with computational engineering's bottom-up (data mining derived) model structures from field datasets available from existing catchment monitoring programs.

This will allow improvement of data mining techniques given expert validation of systems, and also enhance water management by identifying potential bias in our conceptualisation of catchment/water systems relative to the data.

## Expected outcomes

*Highlight the expected outcomes of the project*

- \* New data mining techniques
- \* New catchment and water management models
- \* Analysis and better understanding of potential biases in expert opinions about environmental systems, and the utility of data in moderating this
- \* A PhD student completion
- \* Conference papers
- \* Journal articles: 1. from computational/data-mining perspective, 2. from water management-modelling perspective

## How will the project address the Goals of the above Themes?

*Describe how the project will address the goals of one or more of the 6 Themes listed above.*

This multidisciplinary project directly addresses the “Advanced Computational Engineering, Simulation and Manufacture” theme’s goal of “assessing the veracity of a design principle” and also “addressing the usefulness of Bayesian Networks.

It also addresses the goal of the Water theme for improving the “sustainable management of [...] increasingly strained water resources”.

Developing methods for combining computational (data mining) techniques with expert input will produce better data-mining techniques and more parsimonious, more useful, less biased model structures describing catchment systems for use in water management.

## Capabilities and Degrees Required

*List the ideal set of capabilities that a student should have for this project. Feel free to be as specific or as general as you like. These capabilities will be input into the online application form and students who opt for this project will be required to show that they can demonstrate these capabilities.*

Preferred Capabilities:

- \* Analytical, numerical/quantitative skills and familiarity with statistical methods
- \* Strong communication skills and interest to work on real-world, multi-disciplinary problems
- \* Previous data analysis experience would be a plus