

STATEMENT OF PURPOSE

Baoxiang Pan

Ph.D. Applicant

The unpaved muddy yard after a heavy summer rain is the heaven of earthworms, toads, plants and my childhood's imagination. With the naïve drawing of earning a living without missing the beauty of nature, I entered the Department of Hydrology and Water Resources. The academic trainings during the past 6 years tells that nature will not reveal its true beauty without one's tireless work in decoding it. The academic life is not as cosy and comfortable as I have imagined, but I hold a strong determination to continue this hard way because of the challenges and their aesthetic returns..

Computer codes are my language in expressing academic understandings. The Github time bar (<https://github.com/morepenn>) depicts a fuzzy track of my research interests, which could roughly be divided into two forks, one focusing on hydrological data analysis and the other on hydrological simulation. Experiences tell that well-organized data could reveal mechanism(model) while models could compress data in turn. I am looking for a position to continue my research on connecting and unifying them, which is exactly what the Hydroinformatics and Integrated Hydrology Research Group requires in its program.

As is listed in the CV, I have been writing hydrological models for years. Most of the early codes are too redundant to be reused now. Redundant codes reflect unclear thoughts. Thoughts are in the form of language. Good programmers take themselves as language designers. My endeavour of distilling common structures and making comparisons between specific constitutive functions among various models gradually evolves into a process of constructing a hydrology *Domain Specific Language* (DSL). Different hydrology generations pick up different programming languages (Fortran, C, Matlab,Python, etc.) because of trends, traditions or accidents. I want to pull away their syntactic sugar and extract the hydrology essentials through code abstraction and encapsulation. I believe my knowledge of different programming paradigms (process oriented, object oriented and functional programming) can help a lot in developing the integrated hydrology-agriculture investigation platform of the program.

My strong interest in stochastic hydrology was inspired by the book *Probability, the Logic of Science* (Jaynes E T. ,2003). All scientific knowledge are statistical inferences (C.R.Rao, *Statistics and Truth*), but not all scientists follow this creed in practice. I enjoy the sparkle of wisdom and crystalline rigour in Jaynes's idea of quantifying both scientific induction and theories with probability theory. An important theme in my graduate proposal is to implement this claim in detecting how the water-heat correlation pattern emerges as temporal scale expands. The logical and methodological difficulties forced me to push down and re-construct my fuzzy understanding of probability, stochastic process, information theory and machine learning theory again and again. The effort pays when a seasonal information flow extreme point (represented by conditional entropy) is detected. Data generated from different observation sources and models are constantly flooding in this digital age. One should always stay sober when making stochastic analysis of these data, otherwise he/she would be submerged. I hope my effort in constructing an regulated hydro-agriculture uncertainty estimation platform could help avoid misuse of stochastic methods and improve our understandings of the uncertainties of the system.

Assistant Professor Francisco Muñoz Arriola exchanged his research interests and plans with me during the American Geophysical Union 2014 fall meeting in San Francisco. The congeniality of our talk comes from a similar research interest and same ambition for perfect

code. Francisco has done meaningful work in evaluating the hydrological response to climate and land use changes. The mass workload could be greatly reduced if an integrated agriculture water-state uncertainty estimation platform were constructed as expected. I have the determination and ability in taking an important role in this project.

I got a little thrilled knowing that the other initiator of this program is Professor Dimitri Solomatine. Though have been interested and working on data-driven models for long, I have never expected to be guided by a leader in this area. I used to be obsessed with the specific tricks of iterative evolving in gene algorithm, particle swarm optimizer, SCE-UA algorithm and many other *intelligent* algorithms that could learn *knowledge* from data. The information analysis research I am working on uses support vector regression to distil information from high dimensional hydrological terms. Professor Dimitri has done fundamental work in the fields mentioned above. His understanding of the interface between models and data must be very enlightening.

The United States and Europe are the pioneers of modern civilization. I have long been wishing to integrate into their civilizations which formed the values of the world today. The core of our culture today lies in science. I want to make my contribution to the frontier in my major. For me, the most interesting job is to keep finding secrets of different complex interlocking earth systems. Though sometimes I would complain that the research work turns the beautiful scene into hell once I attempted quantify its material, energy and information flows. But the joy of finding new things pays all. The cross-continental academic experience must be a challenging and exciting journey to make my base in the hydrological academic community. I sincerely hope your favourable consideration regarding my application to the Ph.D. program. Looking forward to joining the big family of HIH!

Sincerely
Baoxiang Pan