YUCHEN WANG

+1 (573) 239-5737   2601 Old Highway 63 South
yuchen.wang@mail.missouri.edu   Columbia, Missouri, 65201

EDUCATION

Master of Arts **in Statistics**, University of Missouri, Columbia, Missouri **expected May 2015**
- Maintained 4.0 cumulative GPA with A+ in Probability Theory, Linear Models and Bayesian Analysis.

Bachelor of Science **in Statistics**, East China Normal University, Shanghai, China **June 2013**
- Took *Coursera* courses CS50, machine learning, high performance computing, etc.
- Nominated for best thesis: *Residual Control Charts based on semi-parametric regression models*

Minor **in Computer Science**, Shanghai Jiao Tong University, Shanghai, China **June 2013**
- Completed all undergraduate curriculum of Computer Science and a graduation thesis.

RELEVENT EXPERIENCE

Research Assistant, University of Missouri, Columbia, Missouri **August 2014 – Present**
- The only research assistant in our graduate class; Collaborated with PhD students in different research projects.
- Optimized a bayesian MCMC algorithm in Fortran for gene data. Reduced computation time from more than 3 hours to less than 20 seconds.
- Parallelized a large-scale Bayes factor computation program in Fortran using both *OpenMP* and *MPI*; Tuned for distributed computing on different kinds of high performance computing clusters.

Research Assistant, East China Normal University, Shanghai, China **July 2012 - June 2014**
- Joined a research team of Statistical Process Control (SPC) in tobacco manufacturing as a part-time research assistant.
  - The only undergraduate student in a research team of 7 members consisting graduate students and faculties.
  - Experienced in analyzing real-world big data (GBs per hour) from observational studies on Amazon EC2 using parallelized algorithms.
  - Designed a control scheme for controlling tobacco dehumidification process, which is the most critical and complicated control unit among all tobacco manufacturing processes.
  - Collaborated with both quality control team at the tobacco manufacturing factory and software engineering team of an IT company. The final quality control product could precisely detect abnormalities of the manufacturing in advance so as to minimize the cost of failure products.
- Hired as a full-time research assistant in charge of the SPC project after graduation. Led a team of 21 members.
  - Mentored 7 members in the team. Trained all new members about previously used models and programs.
  - Developed an R package for the project containing raw data, data cleaning routines, statistical algorithms and visualization tools. Programmed over 2,500 lines of R code for this project.
- Recruited 3 undergraduate students for the development of a statistical software with user interface as a product for a two stage randomized experiment project.
  - Worked remotely with the rest of the team while studying abroad.
  - Designed and developed an interactive data visualization web application in R and *Shiny*.

RELATED SKILLS

Statistical Computing Highly skilled in R; Developer of several R packages (see my GitHub: @wangyuchen); Mostly used packages including *MCMCpack*, *R2OpenBUGS* and *coda* for Bayesian computing; *spatstat* and *splancs* for spatial point pattern analysis; *ggplot2* and *knitr* for graphics and reports. Proficent in SAS procedures including *glm*, *nlin*, *genmod* and *mixed*.

General Purpose Programming Solid understanding of object-oriented programming with *C*++. Extensive parallel computing experience with Fortran (*OpenMP/MPI*). Experience with *Python* and *Java*.

Specialities in Statistics Bayesian analysis; Spatial statistics; Machine Learning and Data Mining; Semi-parametric regression models for longitudinal data and Statistical Process Control.

Mathematical Modeling Proficient in *MATLAB* through 3 Mathematical Contest in Modeling experience. Familiar with operations research techniques in *LINGO*.

Operating Systems Effective in Unix/Linux environment and Shell scripts. Responsible for HPC cluster management at ECNU. Experience in parallel programming on Linux clusters.