

STATEMENT BY CANDIDATE

Statement of Five-Year Goals

SECTION III - B

Instructions: Candidate is to provide a one-page statement of the individual's five-year goals with respect to teaching, research, service, and any other scholarly activity. Statement is limited to the space provided below.

During the next five years, my first goal in research is to organize and write papers for the research done by several of my former students. Most of these students accepted industry jobs and did not have time to work on manuscripts after graduation. I would like to organize the work into manuscripts so that the research results can be published.

My second goal in research is to focus on the new research that my current and future students will be doing with me. One of them is along the lines of Edgeworth expansion in hypothesis testing. One of my current Ph.D. students just worked out several improved tests for two sample comparisons using Edgeworth expansion. Another of my Ph.D. student started to improve the test in high dimensional ANOVA using Edgeworth expansion. So far the results have been promising. I will investigate further into this area and I expect that there is a lot of new research that can be done. A second direction of my research will be in the area of hyperspectral image segmentation, which is needed in satellite image processing. Different from color or grayscale images in natural life, the pixels of such images are in high dimensional feature space. From my past research experience on color or grayscale images as well as high dimensional data analysis, I have come up with an idea to perform hyperspectral image classification and segmentation. This idea will be carried out, implemented, refined, and extensively tested over the next few years.

My third goal in research is to continue collaboration with some entomologists, plant pathologists, and biologists on campus. Even though I do not have appointment with the Agricultural Experimental Station (AES), these collaborators have developed key reliance on me to analyze their data and help write papers to publish their results. I am a critical member to their research. As the most important senior consulting faculty Dr. Murray just retired recently and the AES does not have the resources to fund a senior position, the consulting demand could be hectic to the junior consulting faculty in the department. My collaboration with those researchers on campus can help reduce the burden of the junior consulting faculty.

My goal in teaching is to continue to strive for excellence. I have established a set of methods and principles that generally work well in both classroom teaching and student advising. Over the next five years, I will refine these teaching methods and adjust them based on technology advancement and changes in students. Just as in the past that I have taught a broad spectrum of statistics courses at various levels, I will continue to be a competent, versatile instructor on all statistical courses offered at K-State including applied, theoretical, and computational statistics at both the undergraduate and graduate levels.

My goal in service is to be more actively involved in service to the department, the university and professional society. Both the department and the university need established faculty to play leadership roles. I will embrace this opportunity. Additionally, I expect to spend a significant amount of time providing mentoring for new faculty. We have three junior faculty members that just joined us this Fall semester. One of them is a fresh Ph.D.. Appropriate mentoring can help them avoid pitfalls and reduce frustration. I will offer my advice whenever needed to help them to be productive and feel pleasant working here at KState.

SUMMARY OF CANDIDATE'S INSTRUCTIONAL ACTIVITY

SECTION IV - A

Instructions: Candidate is to provide a one-page summary of courses taught, student advisement, thesis supervision, and any other evidence of instructional productivity. Summary is limited to the space provided below.

During the evaluation period, I **taught 21 courses with topics covering 12 different courses**. Among these courses, one of them is at the undergraduate level for engineering majors (Stat 490), five of them are graduate courses for M.S. students in the statistics major or graduate students in natural sciences or social sciences majors (Stat 702, 703, 716, 745, 770), and the remaining six are graduate level courses for mostly Ph.D. students in the statistics major (Stat 825, 903, 904, 940, 950, 981). Five of the courses (Stat 716, 825, 904, 940, 981) were brand new preparations.

List of courses taught during the evaluation period:

Stat 490 Statistics for Engineering I (Fall 2011, Spring 2013, Fall 2014, Spring 2015)
Stat 702: Statistical Methods for the Social Sciences (Spring 2010)
Stat 703: Statistical Methods for Natural Scientists (Fall 2010, Spring 2013, Summer 2014, Summer 2015)
Stat 716: Nonparametric Statistics (Fall 2011, Fall 2013)
Stat 745: Statistical Graphics (Spring 2010, Spring 2012)
Stat 770: Theory of Statistics (Fall 2013)
Stat 825: Numerical Methods in Statistics (Spring 2011, Spring 2015)
Stat 903: Spatial and Longitudinal Data Analysis (Spring 2011)
Stat 904: Resampling methods (Spring 2014)
Stat 940: Advanced Statistical Methods (Fall 2014)
Stat 950: Data Mining (Fall 2010)
Stat 981: Advanced Inference (Spring 2012)

I also served as the **major professor for 14 graduate students** during the evaluation period. Among them, **four students successfully completed and defended their Ph.D. dissertations and eight students completed their M.S. degree**. Below is the list of students and their defense time (or program of study filing date for current students):

Ph.D. students advised during the evaluation period:

Siti Tolos (defended in Nov 2010); Girly Ramirez (defended in July 2013);
Mohammed Gharaibeh (defended in Summer 2014); Mohammad Sahtout (defended in Summer 2014);
Bo Tong (program of study filed in Fall 2013, to defend in December 2015);
Richard Opoku-Nsiah (Program of study filed in May 2015).

Master students graduated during the evaluation period:

Lei Dong (defended in Dec. 2009); Sharad Silwal (defended in Dec. 2009);
Grace Lee (degree awarded in Dec. 2010); Santosh Ghimire (defended in Spring 2011);
Liang Peng (defended in Summer 2011); Dustin Maurer (defended in Summer 2011);
Eric Mann (defended in Fall 2011); Dayou Jiang (defended in Fall 2012).

Moreover, I served as a **committee member for 15 M.S. and Ph.D. students** and as an **outside chair of three additional Ph.D. committees**. The names are listed here (please see my CV for details):

Committee member for:

Lianqing Zheng, Chun Yu, Zhiwei Sun, Tej Shrestha, Jianjun Hua, Xiuqin Bai, John Richards, Shih-Hsiung Chou, Weilong Cong, Na Qin, Cristiana Rosa Piccinni, Miao Yu, Yaseen Alhaj-Yaseen, Cristina Andreescu, Josh Brummer

Outside chair for Ph.D. committee in Mathematics: Hao Zheng, Jie Ren, Misty Long

SUMMARY OF CANDIDATE'S INSTRUCTIONAL QUALITY

SECTION IV - B

Instructions: Candidate is to provide evidence of instructional quality such as ratings, peer evaluations, evaluation of advisement, outcomes of instructional projects directed, awards, etc. Summary is limited to the space provided below.

A summary of available student ratings on overall effectiveness is listed in the table below.

Course	Semester	credit	#grades assigned	overall effectiveness as a teacher
Stat 702	Spring 2010	3	16	4
Stat 745	Spring 2010	3	22	4.3
Stat 703	Fall 2010	3	44	2.7
Stat 950	Fall 2010	3	7	4.5
Stat 825	Spring 2011	3	9	3.4
Stat 903	Spring 2011	2	8	3.5
Stat 490 section A	Fall 2011	1	48	3.1
Stat 490 section B	Fall 2011	1	42	3.6
Stat 716	Fall 2011	2	9	4.6
Stat 745	Spring 2012	3	39	4.4
Stat 490	Spring 2013	1	68	3.1
Stat 703	Spring 2013	3	46	3.2
Stat 770	Fall 2013	3	12	4.3
Stat 716	Fall 2013	2	17	4.3
Stat 904	Spring 2014	3	2	5.0
Stat 703	Summer 2014	3	15	4.7
Stat 490	Fall 2014	1	65	3.4
Stat 490	Spring 2015	1	70	3.1
Stat 825	Spring 2015	3	7	4.4
Stat 703	Summer 2015	3	12	4.1

Beyond teaching evaluations, another evidence of the quality of my instruction and advisement is that all of my students received job offers (or offers of assistantship for M.S. students to study for Ph.D.) upon graduation. This is a positive recognition that I have been providing high quality advisement for students such that my students are well accepted in industry and academia.

Below is a list of my students' job offers:

Siti Tolos (Assistant Prof., International Islamic Univ. of Malaysia); Girly Ramirez (Novatis, North Carolina); Mohammed Gharaibeh (Assistant Professor, Al Al-Bayt University, Jordan); Mohammad Sahtout (Statistics Specialist, Statistics and Analysis Division, Development Sector, United Arab Emirates); Lei Dong (Statistician in K.U. Medical Center, Department of Internal Medicine); Sharad Silwal (Assistant Professor, Jefferson College of Health Sciences); Santosh Ghimire (Assistant Professor, Tribhuvan University, Kathmandu, Nepal); Liang Peng (Ph.D. student in Computer Science, Utah State Univ.); Dustin Maurer (Ph.D. student in Bioinformatics, University of Kansas); Eric Mann (Consulting Actuary, Streff Insurance Services, Kansas City); Dayou Jiang (Software Engineer, Cerner Corporation, Kansas City); Grace Lee (Ph.D. student in statistics, K-State Univ.).

Bo Tong, a current Ph.D. student of mine already received a job offer in this past summer even before he passed his preliminary exam. These are positive evidences of my instructional quality.

OTHER EVIDENCE OF SCHOLARSHIP AND CREATIVITY IN INSTRUCTION

SECTION IV - C

Instructions: Candidate is to provide any other evidence of scholarship and creativity that promote excellence in teaching such as multimedia presentations, computer-aided instruction, innovative teaching methods, instruction-related publication, presentations, etc. Summary is limited to the space provided below.

Teaching at K-State is an experience of having broad audiences with diverse backgrounds. In the same class, some students were at the Ph.D. level with strong statistical foundations, while some were from other disciplines that could recognize very few statistical terms. To handle such diverse backgrounds and help students learn effectively, I used the following practices: (1) preparing materials well ahead of time and giving students access to all resources before each class, including lecture notes, supplementary reading materials, relevant R programming code, etc. (2) teaching at the level of the audience; (3) engaging the students to learn by adding content step by step and stimulating students through illustrative examples; (4) integrating the latest developments in statistical science into the curriculum; (5) using computer-aided presentations and student hands-on activities (6) providing a positive learning environment and stimulate creative thinking; (7) encouraging students to work hard and develop a positive learning habit.

I have applied these methods in my teaching and advising and have received positive feedback from my students (some of the comments can be seen from TEVALs). In advising, I passed my enthusiasm for statistical research and education on to my students through their interaction with me during advising. Over the period that they worked on their M.S. or Ph.D. research with me, I helped them develop their own competence and passion toward statistical research and applications in applied fields, particularly in the medical sciences. How well their enthusiasm has been nurtured can be seen by their career choice as well as their ability to be hired and thrive with their career choice. Here I list a few consistent examples below.

Dr. Ke Zhang, a former Ph.D. student of mine, was offered a job at Abbott Laboratories to conduct statistical research in Phase I clinical trials more than a year before his graduation. After graduation, he continued to work at Abbott for a year and then accepted a tenure track Assistant Professor position at the School of Medicine and Health Sciences at the University of North Dakota, where he served as the Director of the ND INBRE Bioinformatics Core. He is now an Associate Professor with tenure. After Dr. Zhang left Abbott (current AbbVie), every year AbbVie recruited one student from K-State with my recommendation, either for full time position or internship. This summer AbbVie just offered a full-time position to my current Ph.D. student Bo Tong. Many of my other graduate students also found their enthusiasm in promoting education and scholarship and intellectual growth of statistics upon their graduation: Siti Tolos accepted an Assistant Professor position at the International Islamic Univ. of Malaysia; Mohammed Gharaibeh is an Assistant Professor at Al Al-Bayt University in Jordan. Lei Dong was hired by the University of Kansas Medical Center to provide statistical analysis support for medical researchers. He is now a Senior Research Analyst in the Department of Biostatistics, and the Data Manager for the Office of Scholarly, Academic & Research Mentoring (OSARM), Department of Internal Medicine at University of Kansas Medical Center. Sharad Silwal is now an Assistant Professor of Biomedical sciences at Jefferson College of Health Sciences. Santosh Ghimire is now an Assistant professor of Tribhuvan University, which is the best university in Nepal. These examples tell that I have provided positive impacts on my students' learning or research experience. While working with me, my students have developed a deeper appreciation and passion for statistics through their interactions with me. These are all evidence of my promotion of excellence in teaching and advising.

RESEARCH AND OTHER CREATIVE ACTIVITIES

SECTION V - A

Instructions: Candidate is to provide a one-page statement of research and other creative activities. Statement is limited to the space provided below.

My research during the evaluation period focused on developing effective distribution-free methods for analysis and mining of complex data. In particular, I have been working on and have been producing powerful theoretical and computational methods for nonparametric hypothesis tests in big data, image analysis, high dimensional data modeling and data mining, and applying them to various fields with complex data settings. As a result, I have published a series of articles jointly with my students and coauthors. Two of these articles earned a permanent 'highly accessed' designation from BMC Bioinformatics. Below is a summary of the contributions of my articles on theory or computational methods. They cover three different research areas.

(1) Inference based on original observations or ranks for hypothesis testing in big data. This line of research addressed the asymptotic theory for hypotheses testing in high dimensional ANOVA when the distributions are completely unspecified (Wang and Akritas 2011); provided rank-based inference for testing several effects in nested heteroscedastic functional data that includes a large number of repeated measurements observed within a subject or stratum (Wang and Akritas 2010); proposed a test of independence between the response variable, which can be discrete or continuous, and a continuous covariate after adjusting for heteroscedastic treatment effects (Wang, Tolos and Wang 2010); derived asymptotic procedures as well as finite approximations for the analysis of data arising from series of randomized complete block designs with a large number of factor levels (Bathke et al. 2010); provided a robust nonparametric approach to compare the expressions of longitudinally measured sets of genes under multiple treatments or experimental conditions (Zhang et al. 2011).

(2) Digital image quality assessment and image pixel classification and image segmentation. This line of work was proposed to overcome the drawbacks of popular image similarity measures (mean squared error, signal to noise ratio, structure similarity measure and its variants) that ignore possible nonlinear dependence between the source image and the image being compared. In Wang, Maldonado and Silwal (2011) and Silwal, Wang and Maldonado (2013), we separately proposed nonparametric rank-test-based similarity measures in frequency and wavelet domains. Applications of the methods on a variety of altered images showed superior performance. In Ghimire and Wang (2011), we introduced, implemented and assessed the idea of using combined evidence from the multiple hypothesis testing and minimum distance to carry out image pixel classification and image segmentation. Extensive experiments confirmed that our method has excellent edge detection and texture preservation properties for both grayscale and color images.

(3) Variable selection, model estimation, and data mining in big data. This line of research focused on producing methods for variable selection and modeling in biochemistry, proteomics, and genomics data. The emphasis is high accuracy and being effective for each method applied to many datasets in external validation. In Zhang et al. (2012) and Wang et al. (2013), we introduced two new computational methods for classification of cancer tissue samples based on gene expression data. Each method offers significantly improved leave-one-out classification accuracy across at least 10 data sets compared to popular classifiers. In Qian et al. (2012), Zhou et al. (2012), Li et al. (2012), and Chen et al. (2015), we focused on feature extraction and encoding to summarize genomic sequence information and conduct classification. In Xie et al. (2013) and Dai et al. (2014), we provided pipelines for prediction of multidimensional time series and quantitative structure-activity relationship analysis of peptides. The work introduced methods for both near-neighbor sample selection and feature selection.

Beyond publications, I was invited to give 13 professional talks at national conferences, international conferences, and departmental colloquiums at other universities to present my research.

RESEARCH AND OTHER CREATIVE ACTIVITIES

SECTION V - B

Instructions: Candidate is to provide a list of publications and other creative achievements for the evaluation period. Include items accepted but not yet published/presented.

Peer Reviewed Journal Publications on Theory or Computational Methods

(* marks student authors; † indicates corresponding author; IF refers to journal impact factor)

1. Yuan Chen, Wei Zhou, **Haiyan Wang**, Lanzhi Li, Lifeng Wang, Zheming Yuan† (2015) Prediction of O-glycosylation sites based on Multi-Scale Composition of Amino Acids and Feature selection. *Medical & Biological Engineering & Computing*, Jun;53(6):535-44. doi: 10.1007/s11517-015-1268-9. (IF: 1.726)
2. Lei Jin, Suojin Wang†, and **Haiyan Wang** (2014). A new nonparametric stationarity test of time series in time domain. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. doi: 10.1111/rssb.12091 (IF: 3.515)
3. Zhijun Dai*, Lifeng Wang*, Yuan Chen, **Haiyan Wang**†, Lianyang Bai, and Zheming Yuan† (2014) A pipeline for improved QSAR Analysis of Peptides: physiochemical property parameter selection via BMSF, near-neighbor sample selection via semivariogram, and weighted SVR regression and prediction. *Amino acids*. 46(4):1105-19. doi: 10.1007/s00726-014-1667-5. (IF: 3.293)
4. Yuanguai Xie*, Hongyan Zhang*, **Haiyan Wang**, Lifeng Wang*, and Zheming Yuan† (2013). Prediction of multidimensional time series based on GS-RSR-SVR and its application in agricultural economy. *Bulgarian Journal of Agricultural Science*. 19: 1327-1336. <http://www.agrojournal.org/19/06-24.html>
5. Sharad Silwal*, and **Haiyan Wang**†, Diego Maldonado (2013). Assessment of random-noise contamination in digital images via testing on wavelet coefficients. *Statistics and Its Interface*. 6:117135. DOI: <http://dx.doi.org/10.4310/SII.2013.v6.n1.a11> (IF: 0.702)
6. **Haiyan Wang**†, Hongyan Zhang*, Zhijun Dai*, Ming-shun Chen, and Zheming Yuan (2013). TSG: A new algorithm for binary and multi-class cancer classification and informative genes selection. *BMC Medical Genomics*. 6(Suppl 1):S3. doi:10.1186/1755-8794-6-S1-S3 (IF: 4.073)
7. Hongyan Zhang*, **Haiyan Wang**†, Zhijun Dai*, Ming-shun Chen, and Zheming Yuan† (2012). Improving accuracy for cancer classification with a new algorithm for genes selection. *BMC Bioinformatics*, 13:298. doi:10.1186/1471-2105-13-298 (IF: 2.751)
8. Wei Zhou†, Zhijun Dai*, Yuan Chen, **Haiyan Wang**, Zheming Yuan (2012). High-dimensional descriptor selection and computational QSAR modeling for antitumor activity of ARC-111 analogues based on SVR. *International Journal of Molecular Sciences*, 13, 1161-1172; doi:10.3390/ijms13011161 (IF: 2.464)
9. Santosh Ghimire* and **Haiyan Wang**† (2012). Classification of image pixels based on

minimum distance and hypothesis testing. *Computational Statistics and Data Analysis*, 56, 2273-2287. <http://www.sciencedirect.com/science/article/pii/S0167947312000072> (IF: 1.028)

10. Gang Qian*, **Haiyan Wang**, and Zheming Yuan† (2012). Using homology information from PDB to improve the accuracy of protein β -turn prediction by NetTurnP. *Progress in Biochemistry and Biophysics*. 39(5): 472-482.
http://www.pibb.ac.cn/pibbcn/ch/reader/create_pdf.aspx?file_no=20110370
11. JinLiang Li*, LiFeng Wang*, **Haiyan Wang**, LianYang Bai*, Zheming Yuan†(2012) High-Accuracy Splice Sites Prediction Based on Sequence Component and Position Features. *Genetics and Molecular Research*. 11 (3): 3432-3451.
<http://dx.doi.org/10.4238/2012.September.25.12> (IF: 0.994)
12. Ke Zhang*, **Haiyan Wang**†, Arne C. Bathke, Solomon Harrar, Hans-Peter Piepho, Youping Deng (2011). Gene set analysis for longitudinal gene expression data. *BMC Bioinformatics*. 12:273. Highly Accessed. <http://www.biomedcentral.com/1471-2105/12/273> (IF: 3.028)
13. **Haiyan Wang**†, Diego Maldonado, and Sharad Silwal*(2011). A nonparametric-test-based structural similarity measure for digital images. *Computational Statistics and Data Analysis*. 55:2925-2936. doi:10.1016/j.csda.2011.04.021 (IF: 1.089)
14. **Haiyan Wang**† and Michael Akritas (2011). Asymptotically distribution free tests in heteroscedastic unbalanced high dimensional ANOVA. *Statistica Sinica*. 21(3), 1341-1377. (IF: 0.956)
15. **Haiyan Wang**† and Michael Akritas (2010). Rank test for heteroscedastic functional data. *Journal of Multivariate Analysis*. 101: 1791-1805. (IF: 1.017)
16. Ke Zhang* and **Haiyan Wang**† (2010). Nonparametric tests for longitudinal DNA copy number data. *Statistics and Its Interface*. 3(2), 211-222. (IF: 0.474)
17. Arne C. Bathke†, Solomon W. Harrar, **Haiyan Wang**, Ke Zhang*, and Hans-Peter Piepho (2010). Series of randomized complete block experiments with non-normal data. *Computational Statistics and Data Analysis*. 54(7), 1840-1857. (IF: 1.089)
18. **Haiyan Wang**†, Siti Tolos*, and Suojin Wang (2010). A distribution free nonparametric test to detect dependence between a response variable and covariate in presence of heteroscedastic treatment effects. *The Canadian Journal of Statistics*. 38(3), 408433. (IF: 0.689)

Peer Reviewed Journal Publications on Application or Collaborative Research

19. Zacharakis, Jeffrey, **Wang, Haiyan**, Patterson, Margaret B., & Anderson, Lori (2015). Using modern statistical methods to analyze demographics of Kansas ABE/GED students who transition to community or technical college programs. To appear in *Journal of Research and Practice for Adult Literacy, Secondary, and Basic Education*.

20. Ming-Shun Chen†, Shanda Wheeler, **Haiyan Wang**, R. Jeffrey Whitworth (2014). Impact of temperatures on Hessian Fly (Diptera: Cecidomyiidae) resistance in selected wheat cultivars (Poales: Poaceae) in the great plains region. *Journal of Economic Entomology*, 107(3):1266-1273. <http://www.bioone.org/doi/full/10.1603/EC13357> (IF: 1.605)
21. Sandra Garcs-Carrera, Allen Knutson, **Haiyan Wang**, Kristopher L. Giles, Fangneng Huang, Jeffrey R. Whitworth, Michael C. Smith, and Ming-Shun Chen†(2014) Virulence and biotype analyses of Hessian Fly (Diptera: Cecidomyiidae) populations from Texas, Louisiana, and Oklahoma. *Journal of Economic Entomology*, 107(1):417-423. (IF: 1.605)
22. Chitvan Khajuria, **Haiyan Wang**, Xuming Liu, Shanda Wheeler, John C Reese, Mustapha El Bouhssini, Jeffrey R. Whitworth and Ming-Shun Chen†(2013). Mobilization of lipids and fortification of cell wall and cuticle are important in host defense against Hessian fly. *BMC Genomics*. 14:423. <http://www.biomedcentral.com/1471-2164/14/423> (IF: 4.041)
23. Zenglei Wang, Daniel Parker, Hao Meng, Lanou Wu, Jia Li, Zhen Zhao, Rongping Zhang, Miao Miao, Qi Fan, **Haiyan Wang**, Liwang Cui† and Zhaoqing Yang†, (2012). In vitro sensitivity of plasmodium falciparum from China-Myanmar border area to major ACT drugs and polymorphisms in potential target genes. *PLoS ONE* 7(5): e30927. doi:10.1371/journal.pone.0030927 (IF: 3.73)
24. Lieceng Zhu, Xuming Liu, **Haiyan Wang**, Chitvan Khajuria, John C. Reese, R. Jeff Whitworth, Ruth Welti, and Ming-Shun Chen† (2012). Rapid mobilization of membrane lipids in wheat leaf-sheaths during incompatible interactions with Hessian Fly. *Molecular Plant-Microbe Interactions*. 25(7):920-30. doi: 10.1094/MPMI-01-12-0022-R. (IF: 4.307)
25. Long Cui, Zenglei Wang, Hongying Jiang, **Haiyan Wang**, Xin-zhuan Su, and Liwang Cui† (2012) Lack of association of the S769N mutation in plasmodium falciparum SERCA (PfATP6) with resistance to artemisinins. *Antimicrobial Agents and Chemotherapy*. 56(5):2546-52. <http://www.ncbi.nlm.nih.gov/pubmed/22354307> (IF: 4.565)
26. Xuming Liu, Christie E. Williams, Jill A. Nemacheck, **Haiyan Wang**, Subhashree Subramanyam, Cheng Zheng and Ming-shun Chen† (2010), Reactive oxygen species is involved in plant defense against a Gall Midge. *Plant Physiology*. 152: 985-999. (IF: 6.451)
27. Jung Miao, Qi Fan, Long Cui, Xiaolian Li, **Haiyan Wang**, Gang Ning, Joseph C. Reese, Liwang Cui†(2010) The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite Plasmodium falciparum. *Molecular Microbiology* 78(4): 883902. (IF: 4.819)
28. Shize Zhang, Richard Shukle, Omprakash Mittapalli, Yu Cheng Zhu, John C. Reese, **Haiyan Wang**, Bao-zhen Hua, and Ming-shun Chen† (2010). The gut transcriptome of a gall midge, *Mayetiola destructor*. *Journal of Insect Physiology*. Volume 56, Issue 9,

Peer Reviewed Conference Presentation

29. Zacharakis, Jeffrey & **Wang, Haiyan** (2014, April 5). A pilot study of adult literacy student pathways between secondary and postsecondary education. American Educational Research Association Annual Conference, Philadelphia, PA.

Below is a summary of the impact factor of the journals for my publication as well as the number of times that each article was cited as of August 26, 2015.

Article	impact factor	# of times cited as of Aug. 26, 2015	Year of publication	Article	impact factor	# of times cited as of Aug. 26, 2015	Year of publication
1	1.726		2015	16	0.474		2010
2	3.515	1	2014	17	1.089	6	2010
3	3.293		2014	18	0.689		2010
4			2013	19			2015
5	0.702		2013	20	1.605	1	2014
6	4.073	10	2013	21	1.605	4	2014
7	2.751	24	2012	22	4.041	6	2013
8	2.464	10	2012	23	3.73	19	2012
9	1.028	3	2012	24	4.307	8	2012
10			2012	25	4.565	24	2012
11	0.994	7	2012	26	6.451	70	2010
12	3.028	6	2011	27	4.819	13	2010
13	1.089	4	2011	28	2.31	13	2010
14	0.956	6	2011	29			2014
15	1.017	4	2010				

Beyond peer reviewed publications and refereed conference presentation, I also gave 13 invited talks at national and international conferences and departmental seminars in other universities. In addition, I co-organized an invited session to promote interaction of young researchers with senior researchers in a conference in Austria in 2015 even though I could not attend the conference myself. Here is the list of the talks I gave or organized during the evaluation period in reverse chronological order:

1. Co-organizer of an invited session “Modeling of Complex Biological Data” for ISNPS 2015 Conference on “Biosciences, Medicine, and novel Non-Parametric Methods”, 12th-15th of July, 2015, in Graz, Austria.
2. Feature selection, sample selection, and prediction for quantitative structure-activity relationship analysis of peptides. Invited talk at 6th International Statistics Forum, May 24-25, 2014, Renmin University, Beijing, China.
3. Improved Prediction Analysis of Microarray (PAM) through different thresholding

- methods and Heteroscedastic Modeling. Invited talk at The Fourth International Biostatistics Workshop, June 4-8, 2014, Jilin University, Changchun, China.
4. A QSAR analysis of peptides: physiochemical property parameter selection, near-neighbor sample selection, and weighted SVR regression and prediction. Invited talk at Second Conference of the International Society for NonParametric Statistics. Cadiz, Spain. June 12-16, 2014.
 5. An algorithm for binary and multi-class cancer classification and informative genes selection. Contributed talk at 2013 Joint Statistical Meeting, Montreal, Canada, July 2013.
 6. Distribution-free Tests for Unbalanced Heteroscedastic longitudinal Data in high dimensional ANOVA settings. Invited talk at 2012 Joint Statistical Meeting, San Diego, July 2012.
 7. Nonparametric Variable Selection in High Dimensional Data (with continuous response). Invited talk at 1st Conference of International Society of NonParametric Statistics, Greece, June 2012.
 8. Nonparametric variable selection in high dimensional data for classification. Invited seminar talk at Department of Statistics, Texas A&M university, March 22, 2012.
 9. A Distribution-Free Test to Detect General Dependence and High Dimensional Variable Selection. Invited talk in special session of 'Mathematical Statistics' at the 1081st AMS Meeting. March 30 - April 1, 2012.
 10. Nonparametric Variable Selection in High Dimensional Data for Classification. Invited seminar talk at Department of Statistics, Penn State University. Aug. 2011.
 11. Nonparametric tests and test-based genome partitioning for longitudinal DNA copy number data. Invited seminar talk at Department of Mathematics and Statistics, University of Missouri at Kansas City. Oct. 2010.
 12. K-nearest neighbor based distribution-free tests to identify relationship between variables in presence of potentially a large number of categorical factors. Invited seminar talk at Department of Statistics, University of Missouri. Sep. 2010.
 13. A new image similarity index based on nonparametric hypothesis testing. Invited talk at the International Conference on Statistical Analysis of Complex Data. Kunming, China, July 2010.
 14. A nonparametric test of nonlinear covariate effects and their interactions with categorical factors. Invited seminar talk at Department of Statistics, Shanghai Fiance and Economics University. July 2010.
 15. A Nonparametric Test of Independence and Its Application in Digital Image Quality Assessment, Invited Project PACE speaker for Colloquium at Department of Mathematical Sciences, University of Montana (PACE represents Partnership for Comprehensive Equity with the goal to increase the representation and advancement of women in academic science and engineering careers). May 2-6, 2010.

RESEARCH AND OTHER CREATIVE ACTIVITIES

SECTION V - C

Instructions: Candidate is to provide a list of grants and contracts funded during the evaluation period. Include agency, funding level, duration, title, and collaborators. Candidate may provide a separate list of grants and contracts applied for, but not funded during the evaluation period.

Award and Funded grants, contracts, or other funded creative activities:

1. 2010 President's Faculty Development Award. This award provided \$2500 to support my travel to 2010 International Conference on Statistical Analysis of Complex Data (in Kunming, China) to give an invited talk.
2. Funding Agency: Simons Foundation
Project Title: Collaboration Grant Proposal on Methods for High Dimensional Data
Award Number: 246077
Award period: 09/1/12 to 08/31/17
Role: PI Total amount: \$35000
Effort: Single author grant.
3. Funding agency: Kansas Board of Regents (Kansas Education Research Scholars Program)
Project title: Exploration of KBOR and KSDE data to better understand Participation, retention and graduation patterns of ABE/GED students transitioning to community college programs
Funding period: 03/25/2013 to 08/01/2014
PI: Jeffrey Zacharakis
Role of Haiyan Wang: Collaborator
total: \$10,500 (among which 0.2185 FTE x 1 month summer salary is for Haiyan Wang)

Grant applications submitted but not funded.

1. Title: Using State Longitudinal Data to Identify Postsecondary Education Pathways of Adult Education Students. Role: Co-PI. (PI: Zacharakis, Jeffrey). Total requested amount: \$371,453
2. Title: Kansas Assessment of Adult learners Transitioning to Postsecondary STEM Programs. Total requested: \$249,911 for 07/01/2011- 06/30/2013. PIs: Jeffrey T Zacharakis, Jacqueline D Spears, Haiyan Wang. This proposed activities are mainly statistical analysis of public data using generalized additive (or linear) mixed models to identify key variables affecting the probability adult learners transitioning to STEM programs. Funding agency: NSF.
3. Title: Kansas Assessment of Adult Learning Practices (KAALP): Utilizing State-wide Adult Education Data to Determine Successful Reading, Writing, and Mathematics Instructional Practices and Instructional Intensity for Struggling Adults. Role: Co-PI PI: Jeffrey Zacharakis. Other Co-PIs: Tim Frey, Weixing Song. Funding agency: U.S. Department of Education.
4. Title: BIGDATA: Mid-Scale: ESCE: DA: Collaborative Research: SciDataCloud - A Data and Collaboration Platform for Advancing Science and Engineering. Requested support period 1/1/2013-12/31/2017; requested amount: \$389,533. Role: Co-PI. Funding agency: NSF.
5. Title: BIGDATA: Small: DA: New Data Mining Tools for Feature Selection and Model Prediction. Requested support period 7/1/2013 -6/30/2016; requested amount \$624,470. Role: PI. Funding agency: NSF.
6. Title: BioQ: A Cloud-based Web System for Enhancing Quantitative Skills of Bioscientists. Requested support period 4/1/2013-3/31/2017; amount \$116,842. Role: Co-PI. Funding agency: NIH.
7. Title: Using State Data to Identify Pathways and Pipelines of Adult Literacy Students. Requested amount: \$400,000. Duration: 2 Years. Role: Collaborator. PI: Jeffrey Zacharakis. Requested funding period: 7/1/2014 to 6/30/2016. Funding agency: US Department of Education (Institute of Education Sciences)

SERVICE CONTRIBUTIONS

SECTION VI

Instructions: Candidate is to provide a statement of service contributions during the evaluation period. Statement should provide evidence of leadership. A list of committees on which the person served may be provided. Statement and committee listing may not exceed two pages.

During the evaluation period, I was actively involved in services at the department level, university level, and professional societies. My leadership roles at the department level include chair of doctoral assessment committee, seminar chair, coordinator of qualifying exam committees, and chair for Linear Models Exam subcommittee. For service to the university, I have been serving as a member in two of the Faculty Senate Committees; I am also a member of the Presidential Committee on Big 12 Faculty Fellowship; I am a Graduate Council representative for Mathematical and Engineering Sciences Academic Area; and I was elected to be the Chair for Graduate Council Academic Affairs Committee. For service to professional society, I have served as an editorial board member for a few journals since 2011 (Associate Editor for Journal of Nonparametric Statistics, review editor for Frontiers in Bioinformatics and Computational Biology, Associate Editor for Statistics Research Letters, Open Journal of Statistics, Austin Statistics) and became an elected member of the International Statistical Institute (ISI) in 2012. I served as a reviewer to two tenure and/or promotions cases in other institutions. I also served as a reviewer for manuscripts submitted to numerous journals and a book reviewer for several publishers. Below is a more detailed list of my services:

Service to the Department

Chair of Doctoral assessment committee (Fall 2014 - Summer 2015)
Coordinator of Ph.D. qualifying exam committees (Fall 2014 -Summer 2015)
Ph.D. qualifying exam committee for Linear Models Exam (member during Jan. to July 2013, chair during Aug - Dec. 2013)
MS Exam on Theory (Aug. 2013 to Summer 2015)
Promotion, tenure and evaluation document revision committee (Fall 2010)
Applied Statistics in Agriculture Conference, Chair for local arrangement (Spring 2010 -Fall 2011)
Masters assessment committee (Fall 2011 - Aug. 2012)
Seminar Chair (Aug. 2011 - Dec. 2011, Fall 2014, Spring 2015)
Faculty Search Committee (Fall 2010)
Statistics in Agriculture - local arrangements (Spring 2010)
Departmental scholarships and awards (Spring 2010)
Library liaison (Spring 2010 - Aug 2013)

Service to the University

- Faculty Senate Committees (Aug. 2014- 2017)
 - Faculty Salaries and Fringe Benefits committee
 - University Library Committee
- Presidential Committee on Big 12 Faculty Fellowship (Fall 2014 - 2017)
- Graduate Council representative for Mathematical and Engineering Sciences Academic Area (September 2015 - August 2018)
- Chair for Graduate Council Academic Affairs Committee (Aug. 2015 – July 2016)

Service to the professional society

Jan. 2012 - current, Elected member of the International Statistical Institute (ISI).
May 2014 - current, Associate Editor for Austin Statistics.
July 2012 - current, Associate Editor for Statistics Research Letters.
Jan. 2011 - current, Associate Editor for Journal of Nonparametric Statistics.

Dec. 2011 - current, Associate Editor for Open Journal of Statistics.

April 2011 - current, Review Editor for Frontiers in Bioinformatics and Computational Biology.

Served as a member for International Biometric Society Young Statistician Showcase Committee. In charge of selecting the Best Student Paper Award from all applicants in the continent of America attending IBC 2012 in Kobe, Japan.

Served as a reviewer for a tenure & promotion case for a faculty of Department of Mathematics and Statistics at Old Dominion University.

Served as a reviewer for a promotion case for a faculty of the Department of Information Sciences at City of Hope National Medical Center.

Served as a reviewer for Journal of the American Statistical Association, Statistica Sinica, Annals of Applied Statistics, Journal of Nonparametric Statistics, Statistica Neerlandica, Journal of Classification, Journal of Royal Statistical Society (serial B), Journal of Systems Science and Complexity, Computational Statistics and Data Analysis, Statistics Research Letters, Journal of the Chemical Society of Pakistan, BMC Medical Genomics, BMC Bioinformatics, BMC Genomics, Computational Statistics and Data Analysis, Frontiers In Bioinformatics and Computational Biology, British Journal of Mathematics & Computer Science, Annals of the Institute of Statistical Mathematics, Statistical Papers, Bioinformatics.

Served as a book reviewer for Prentice Hall, JohnWiley & Sons, Ltd., Psychology Press/Routledge, Taylor and Francis, Biometrics, and Journal of American Statistical Association/ The American Statistician (JASA/TAS).