Multidimensional Analysis of Snow Cover Data in South Dakota Diversity of Landscapes

Photo picture taken:
Thursday, March 27, 2008,
7:45 am

Shmagin, B.A. & C.A. Johnston
Water Research Institute Department of Biology & Microbiology

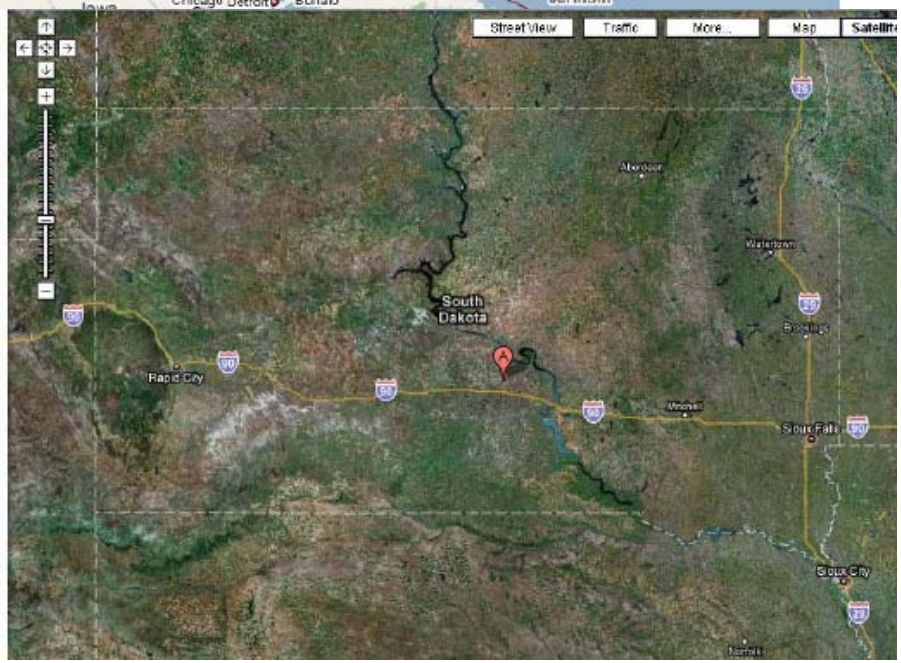
South Dakota State University,
Brookings, SD

Topics

- Introduction: maps & snow fall as the most reliable characteristic
- Data & philosophy of analysis
- Analysis of snow fall for period 1955-79
- Regime of snow fall in SD
- Improving the knowledge & creating maps of natural resources of SD
- Questions

Maps as a universal tool

Google Maps

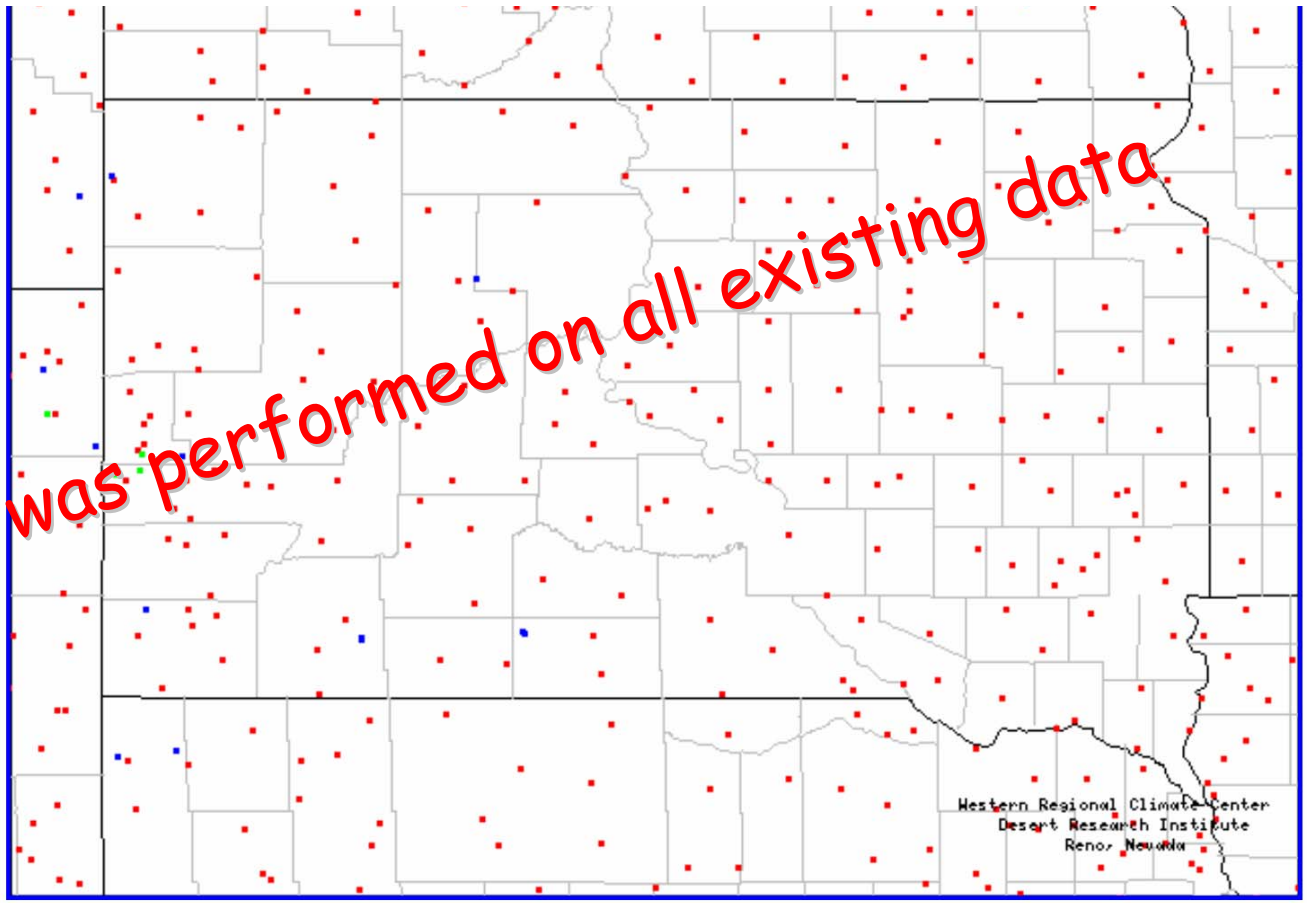


http://hprcc.unl.edu/cgi-bin/cli_perl_lib/cliMAIN.pl?sd1076

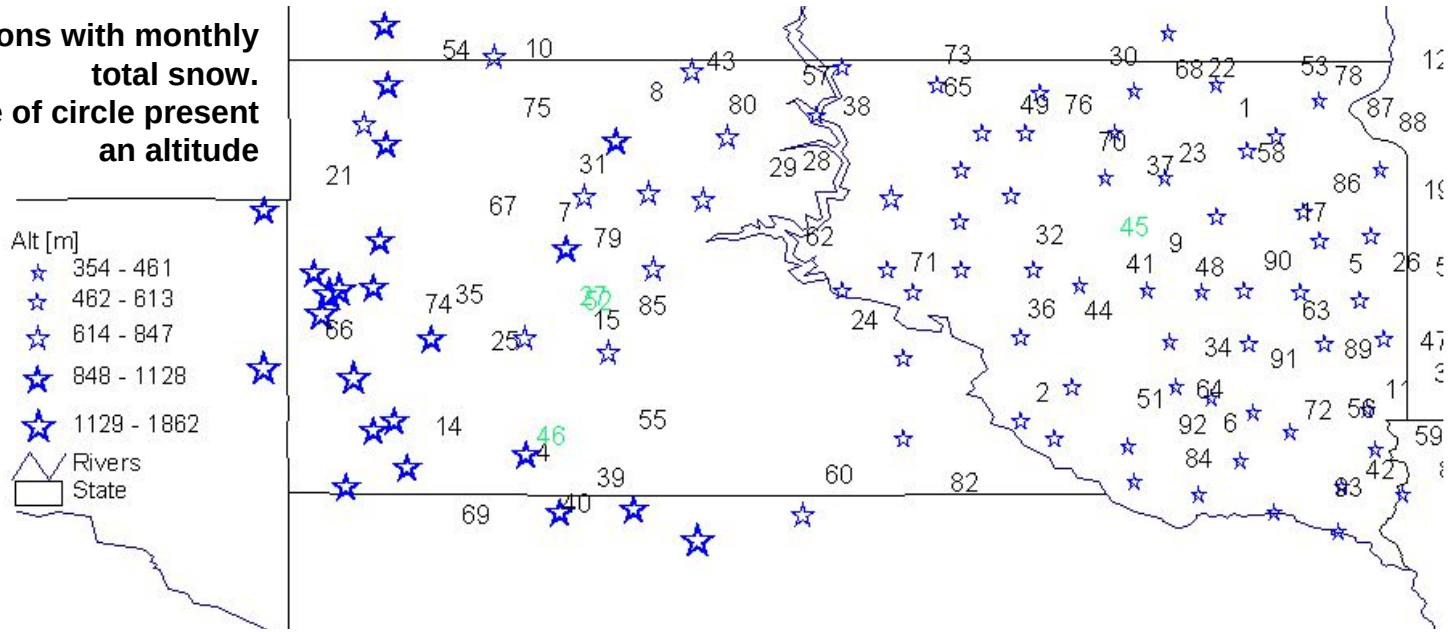
Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Data & philosophy of analysis

The analysis was performed on all existing data



Data on monthly snow fall in SD



Initial matrix: $X_{(n*p)}$ or $X_{(93*11)}$ there are:

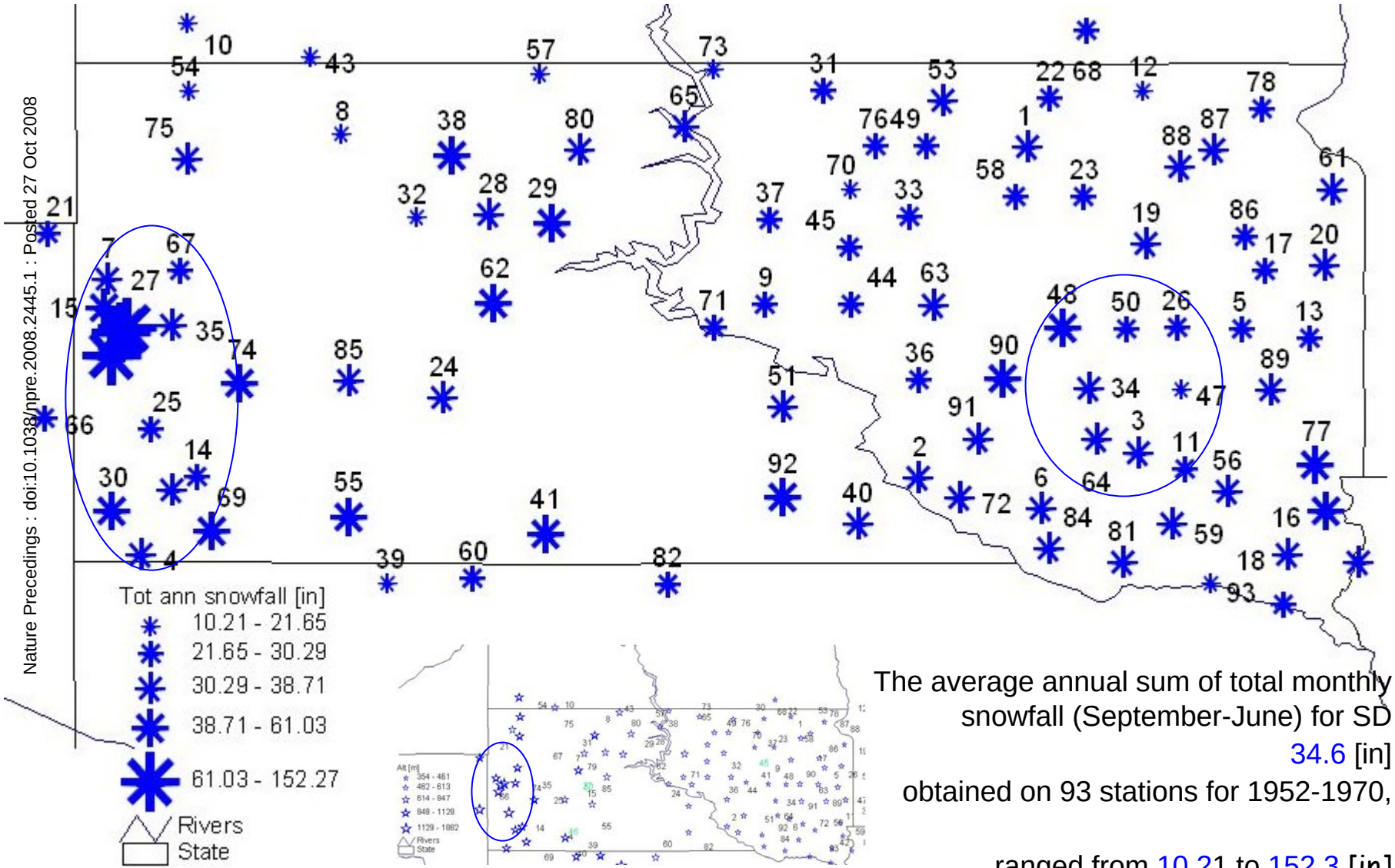
n=93 – number of rows or stations, those stations have mutual time interval of 18 winters of observations (1952-53 – 1969-70),

p=11 – number of variables or 10 monthly proportions & annual total snowfall for the winter season

Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Average annual snow fall in SD

Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

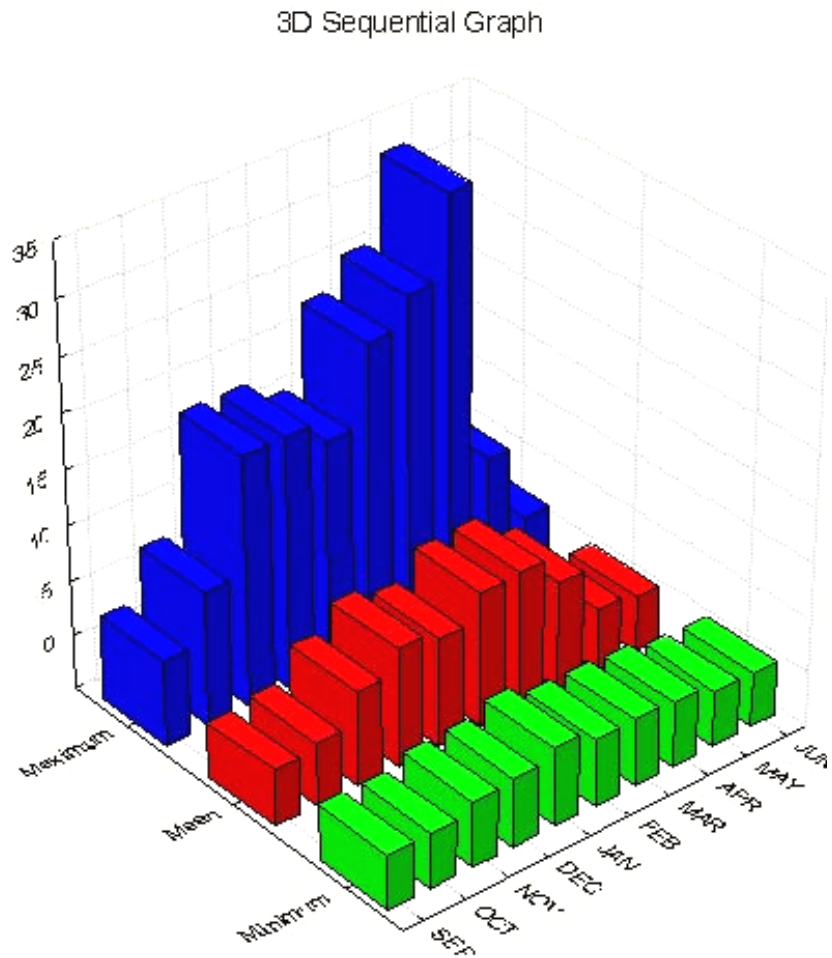


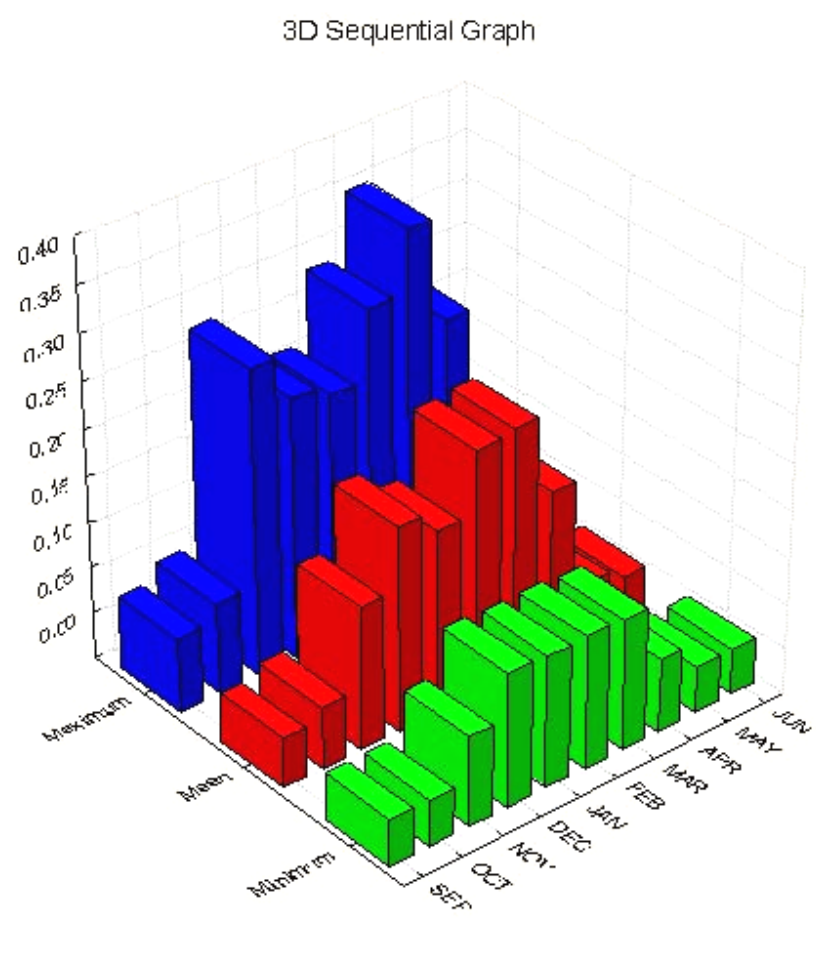
Stations with monthly total snow.
Size of circle present an altitude

The average annual sum of total monthly snowfall (September-June) for SD 34.6 [in]
obtained on 93 stations for 1952-1970,

ranged from 10.21 to 152.3 [in]

Average annual snow fall in SD





The average total monthly snowfall [in] obtained on 93 stations for 1952-1970 (left) & monthly proportion of total annual

Data & philosophy of analysis

Multivariate analysis is the simultaneous statistical consideration of relationships among many measured properties of a given system (Gould 1996, p. 42;).

A **factor** is a portion of a quantity, usually an integer or polynomial that, when multiplied by other factors, gives the entire quantity.

The determination of factors is called factorization (or sometimes "factoring"). It is usually desired to break factors down into the smallest possible pieces so that no factor is itself factorable.

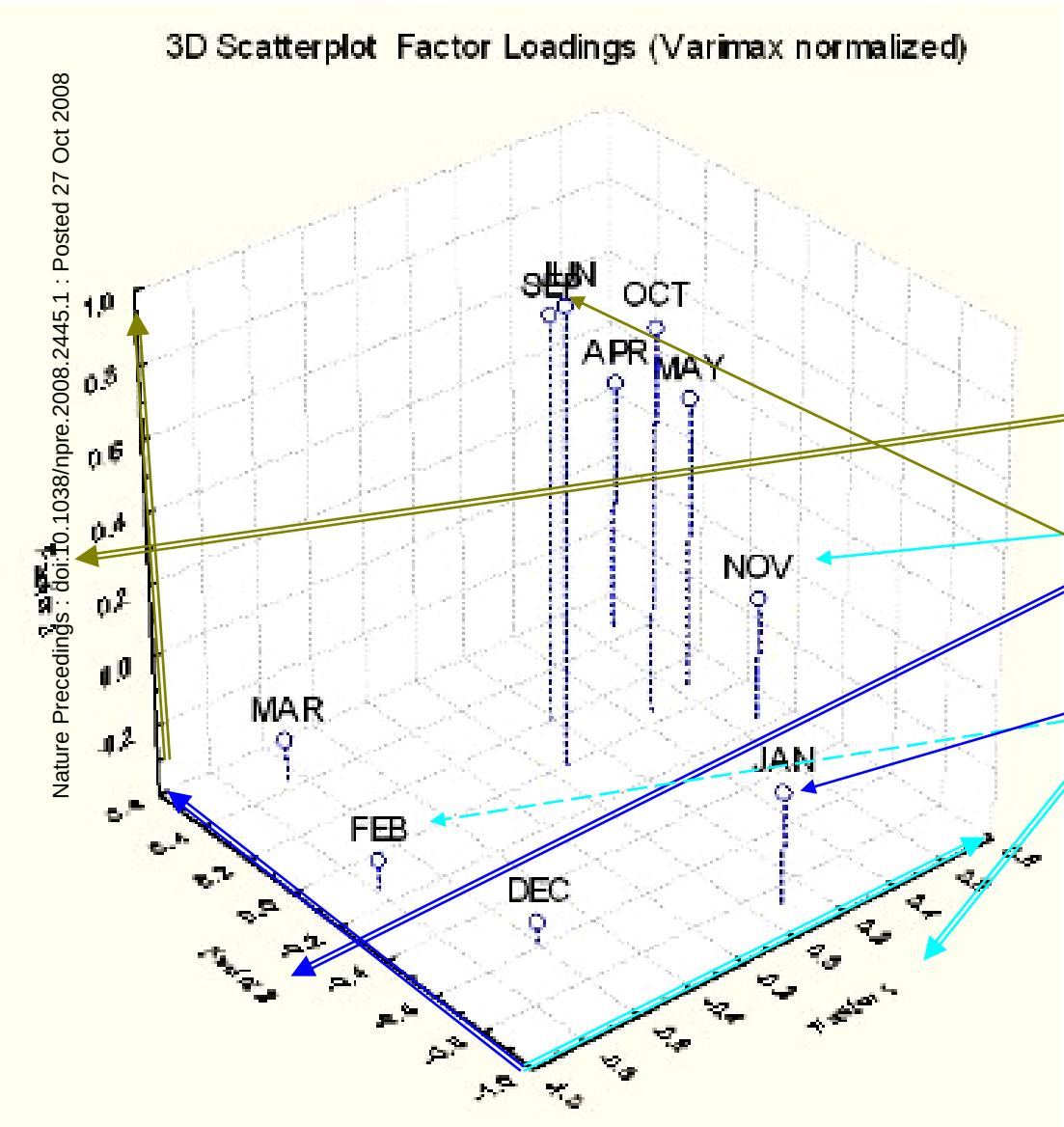
Factor analysis allows the determination of common axes influencing sets of independent measured sets.

The main applications of factor analytic techniques are:

- (1) to *reduce* the number of variables &
- (2) to *detect structure* in the relationships between variables.

(From: Wolfram *MathWorld*)

Structure of monthly snow fall

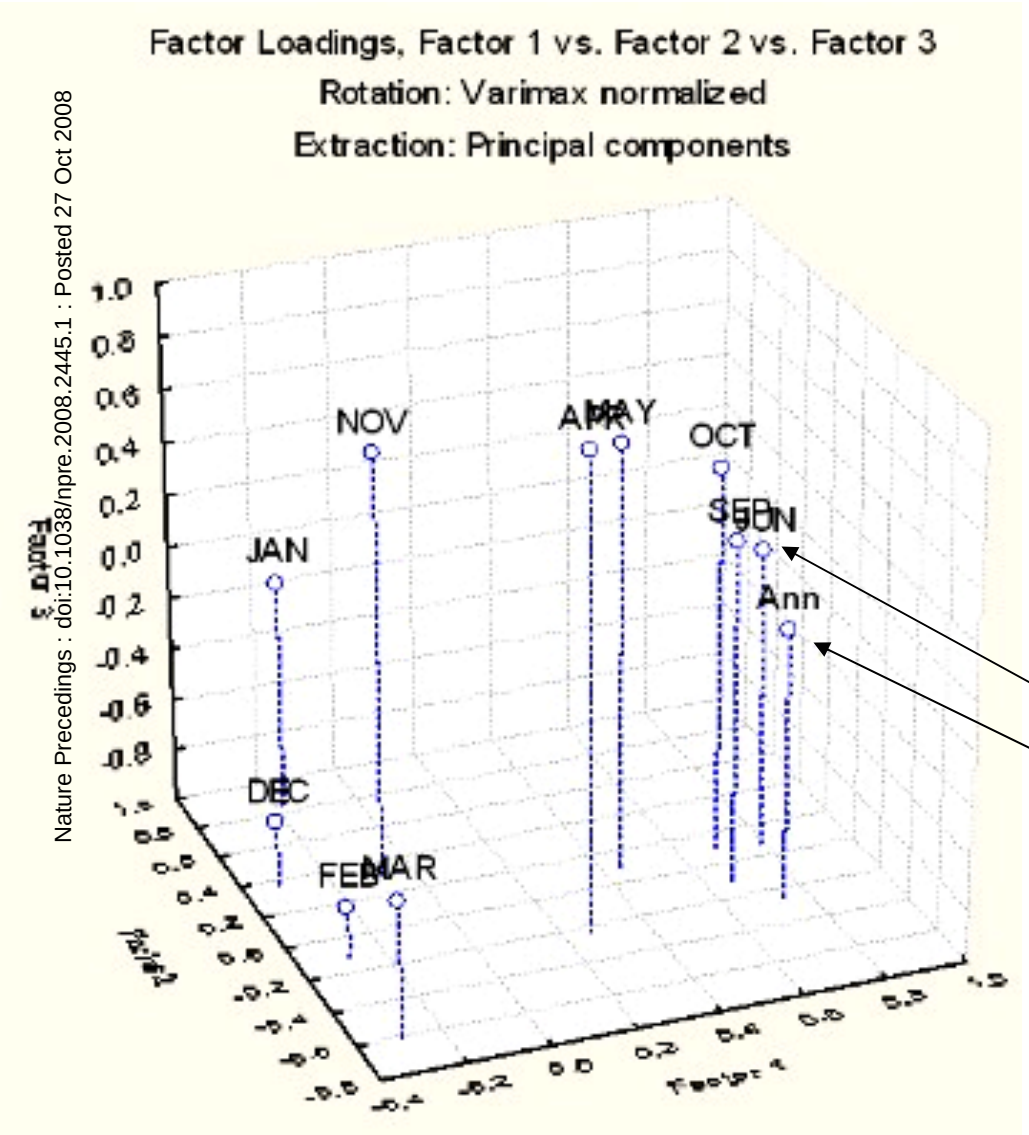


Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Factor Loading

	Factor 1	Factor 2	Factor 3
SEP			**0.74**
OCT	0.46		***0.69***
NOV	***0.69***		
DEC	***-0.59***	***-0.58***	-0.35
JAN		**-0.86**	
FEB	**-0.79**		-0.33
MAR	***-0.65***	0.43	-0.29
APR	***0.69***	0.46	0.31
MAY	***0.65***		0.42
JUN			**0.88**
Expl.Var	3.03	1.58	2.40
Prp.Totl	0.30	0.16	0.24

Annual & monthly snow fall in SD



Factor Loading

	Factor 1	Factor 2	Factor 3
SEP	***0.66***		0.31
OCT	***0.69***		0.47
NOV		0.32	***0.62***
DEC	-0.31	0.37	**-0.76**
JAN		**0.85**	
FEB	-0.28		**-0.81**
MAR	-0.30	***-0.60***	-0.48
APR	0.27		**0.82**
MAY	0.45		***0.63***
JUN	**0.79**		
Ann	**0.74**	-0.25	
Expl.Var	2.66	1.49	3.26
Prp.Totl	0.24	0.14	0.30

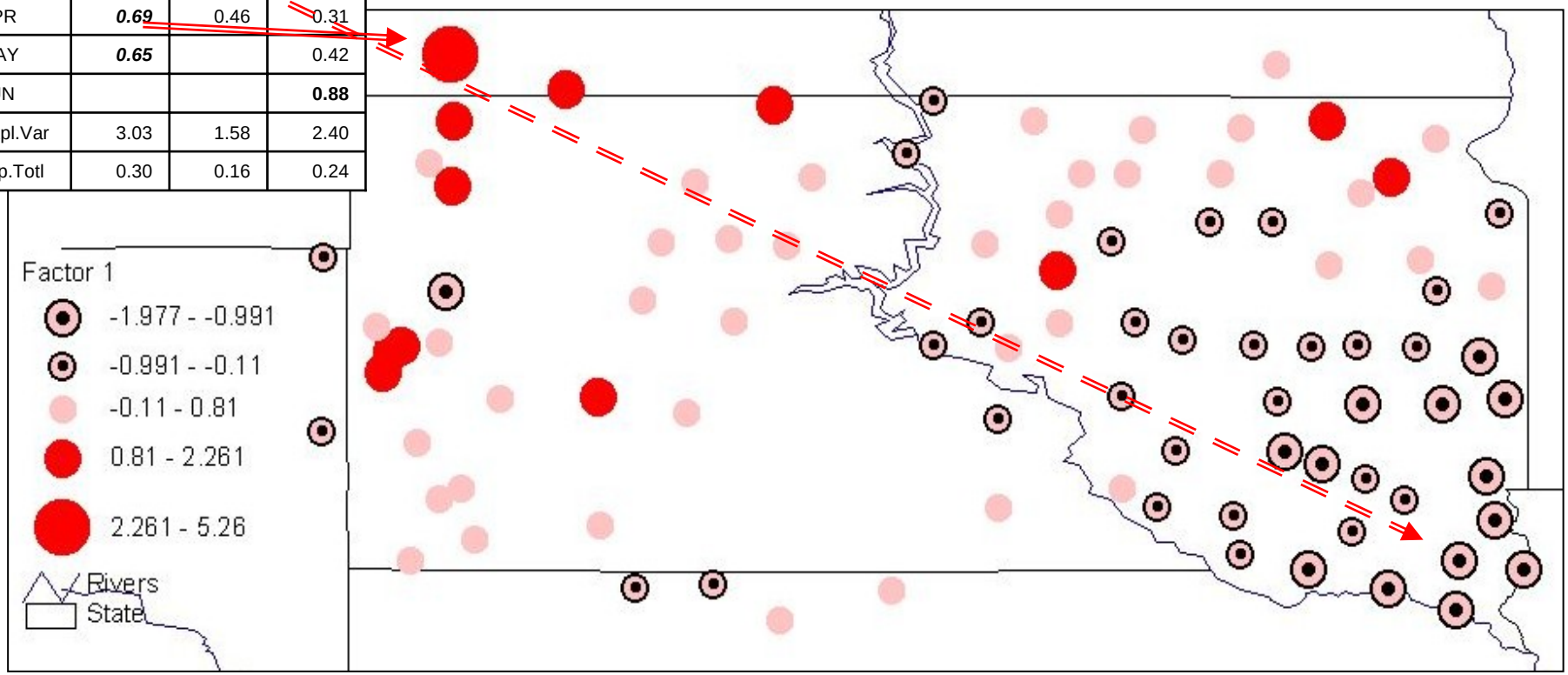
Distribution of monthly snow fall

	Factor 1	Factor 2	Factor 3
SEP			**0.74**
OCT	0.46		***0.69***
NOV	***0.69***		
DEC	***-0.59***	***-0.58***	-0.35
JAN		**-0.86**	
FEB	**-0.79**		-0.33
MAR	***-0.65***	0.43	-0.29
APR	***0.69***	0.46	0.31
MAY	***0.65***		0.42
JUN			**0.88**
Expl.Var	3.03	1.58	2.40
Prp.Totl	0.30	0.16	0.24

Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Factor 1 Scores distribution

Size of circle present Factor Scores



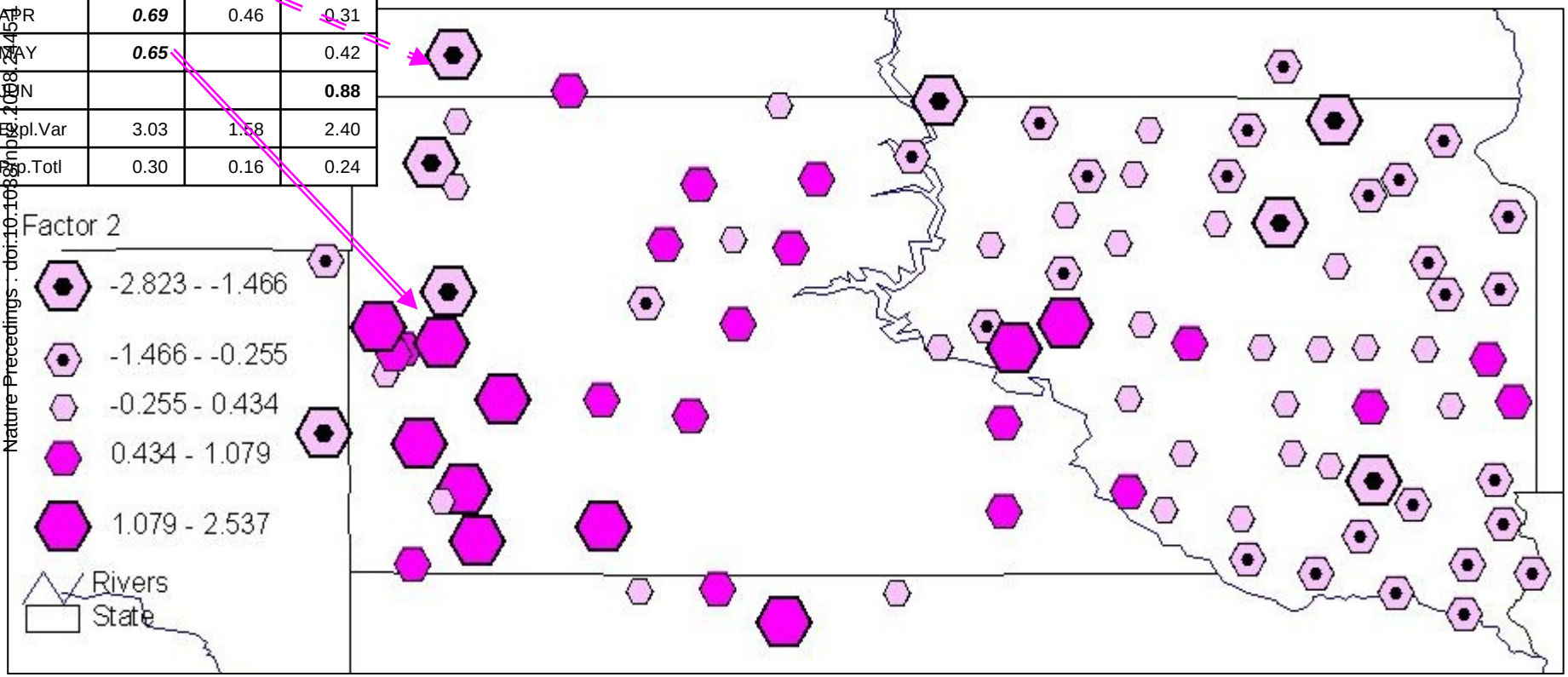
Distribution of monthly snow fall

	Factor 1	Factor 2	Factor 3
SEP			**0.74**
OCT	0.46		***0.69***
NOV	***0.69***		
DEC	***-0.59***	***-0.58***	-0.35
JAN		**-0.86**	
FEB	**-0.79**		-0.33
MAR	***-0.65***	0.43	-0.29
APR	***0.69***	0.46	0.31
MAY	***0.65***		0.42
JUN			**0.88**
Expl.Var	3.03	1.58	2.40
Prp.Totl	0.30	0.16	0.24

Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 2 Oct 2008

Factor 2 Scores distribution

Size of circle present Factor Scores

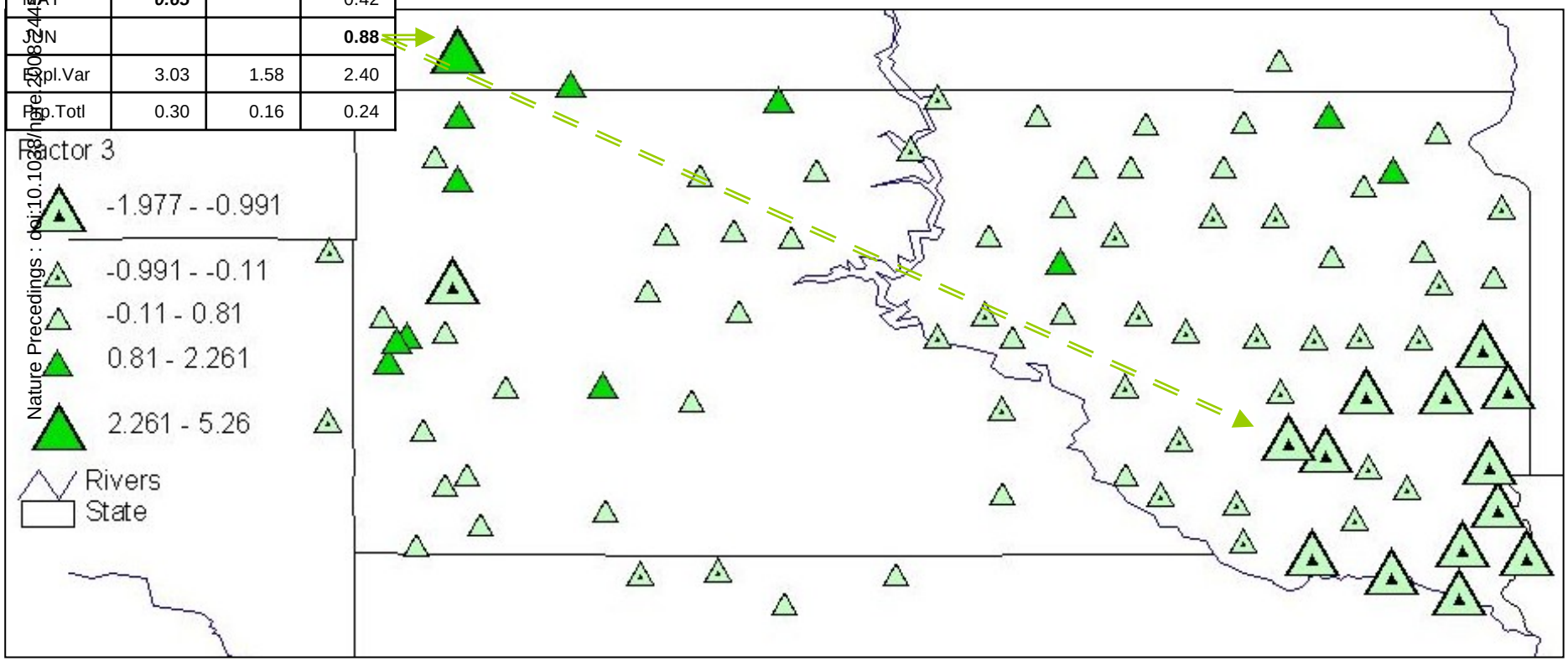


Distribution of monthly snow fall

	Factor 1	Factor 2	Factor 3
SEP			**0.74**
OCT	0.46		***0.69***
NOV	***0.69***		
DEC	***-0.59***	***-0.58***	-0.35
JAN		**-0.86**	
FEB	**-0.79**		-0.33
MAR	***-0.65***	0.43	-0.29
APR	***0.69***	0.46	0.31
MAY	***0.65***		0.42
JUN			**0.88**
Expl.Var	3.03	1.58	2.40
Prp.Totl	0.30	0.16	0.24

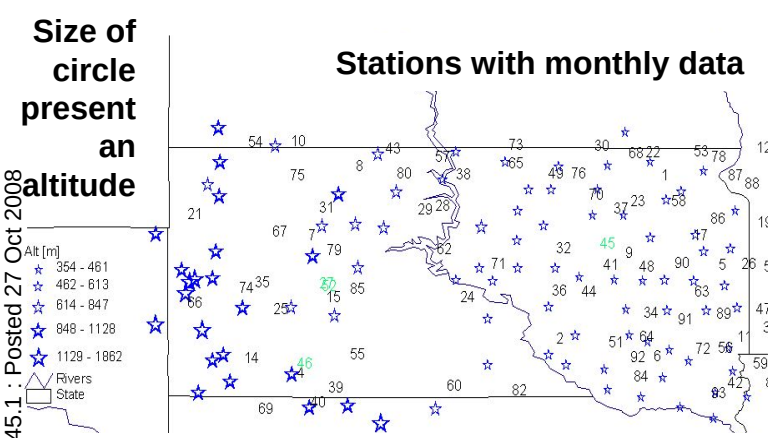
Factor 3 Scores distribution

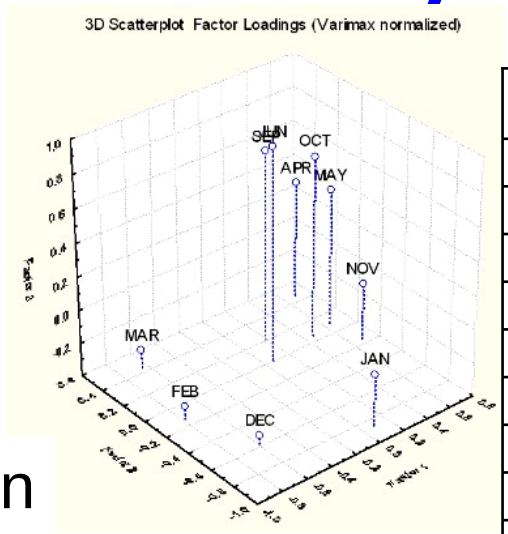
Size of circle present Factor Scores



Nature Precedings : doi:10.1038/npre.2008.2441.1 : Posted 17 Oct 2008

Structure of monthly snow fall





	[in]	Factor 1	Factor 2	Factor 3
SEP	.15			**0.74**
OCT	.66	0.46		***0.69***
NOV	3.6	***0.69***		
DEC	5.7	***-0.59***	***-0.58***	-0.35
JAN	4.9		**-0.86**	
FEB	7.2	**-0.79**		-0.33
MAR	7.6	***-0.65***	0.43	-0.29
APR	4.9	***0.69***	0.46	0.31
MAY	.57	***0.65***		0.42
JUN	.04			**0.88**
Expl.Var		3.03	1.58	2.40
Prp.Totl		0.30	0.16	0.24

Factor Scores distribution

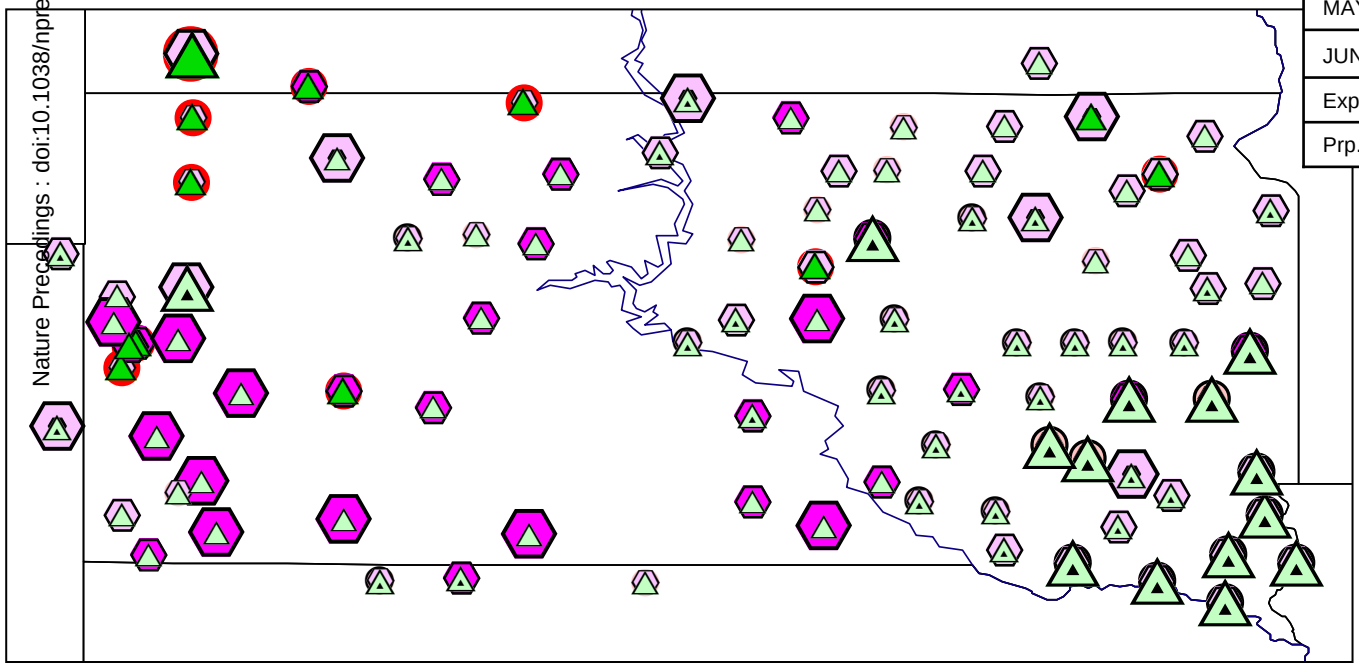
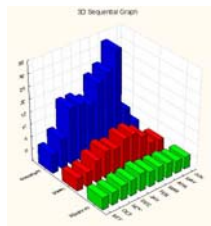
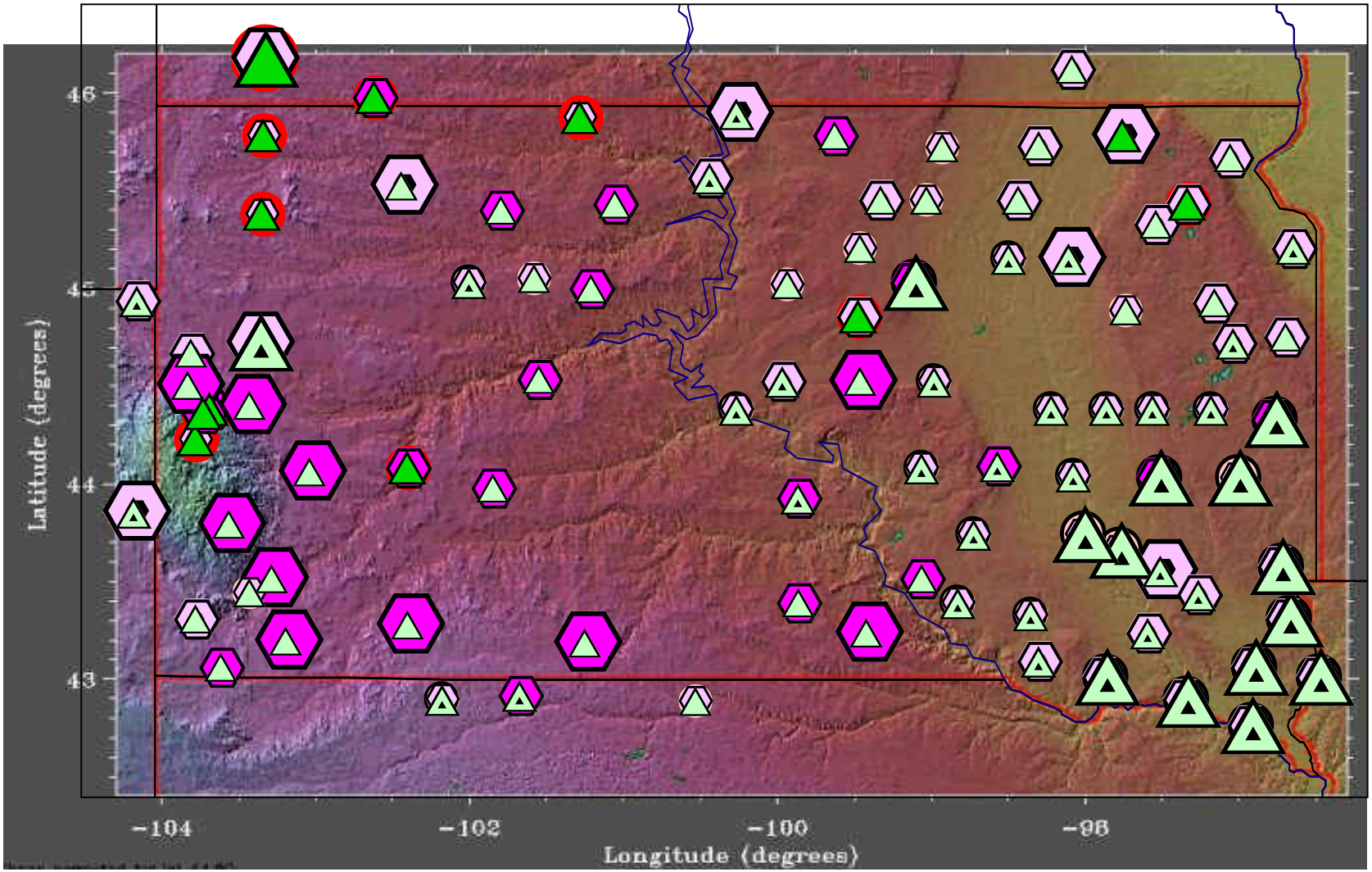


Table of average total annual snow fall & Factor Loading



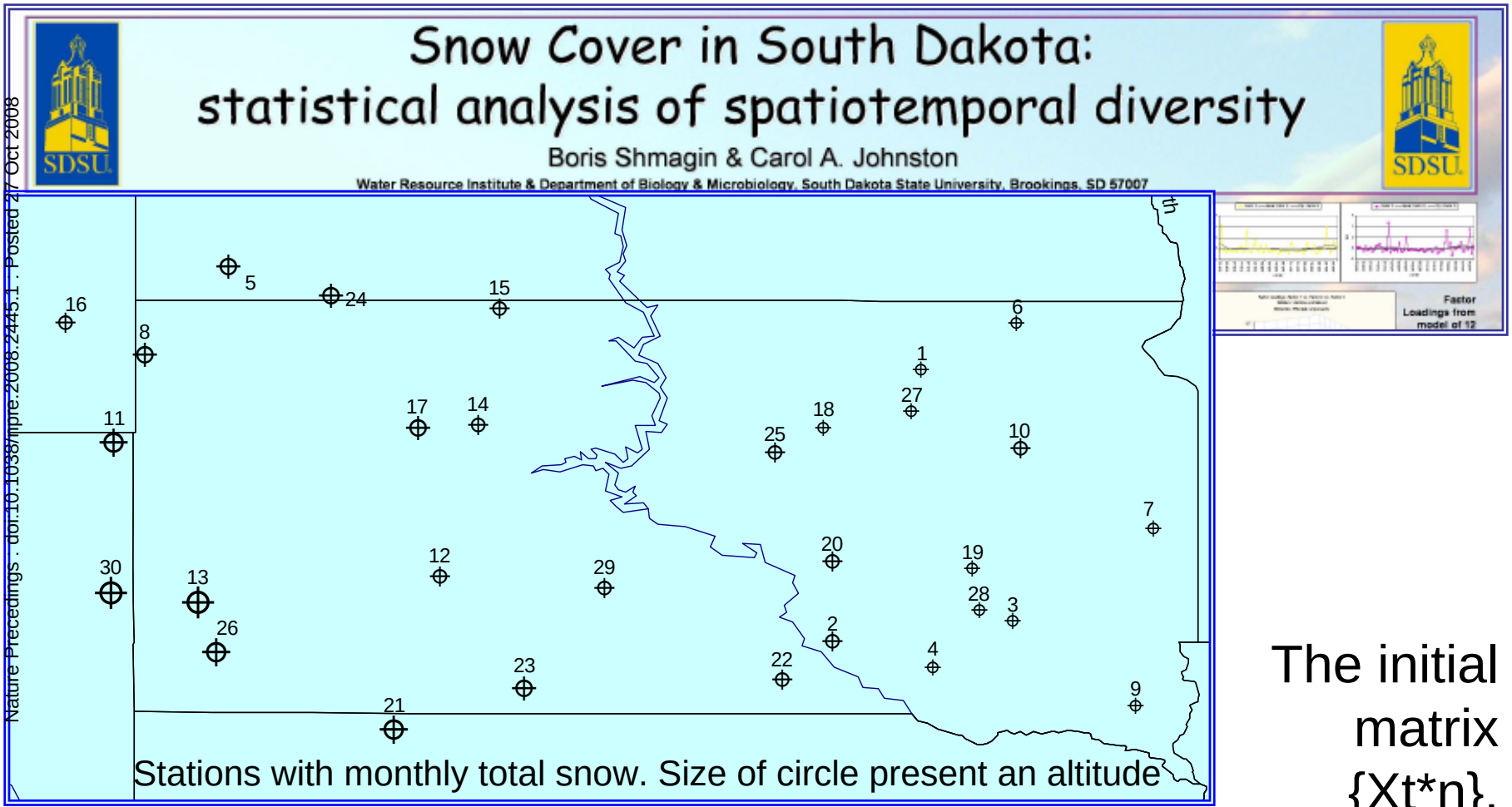
Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Snow fall & landscapes of SD



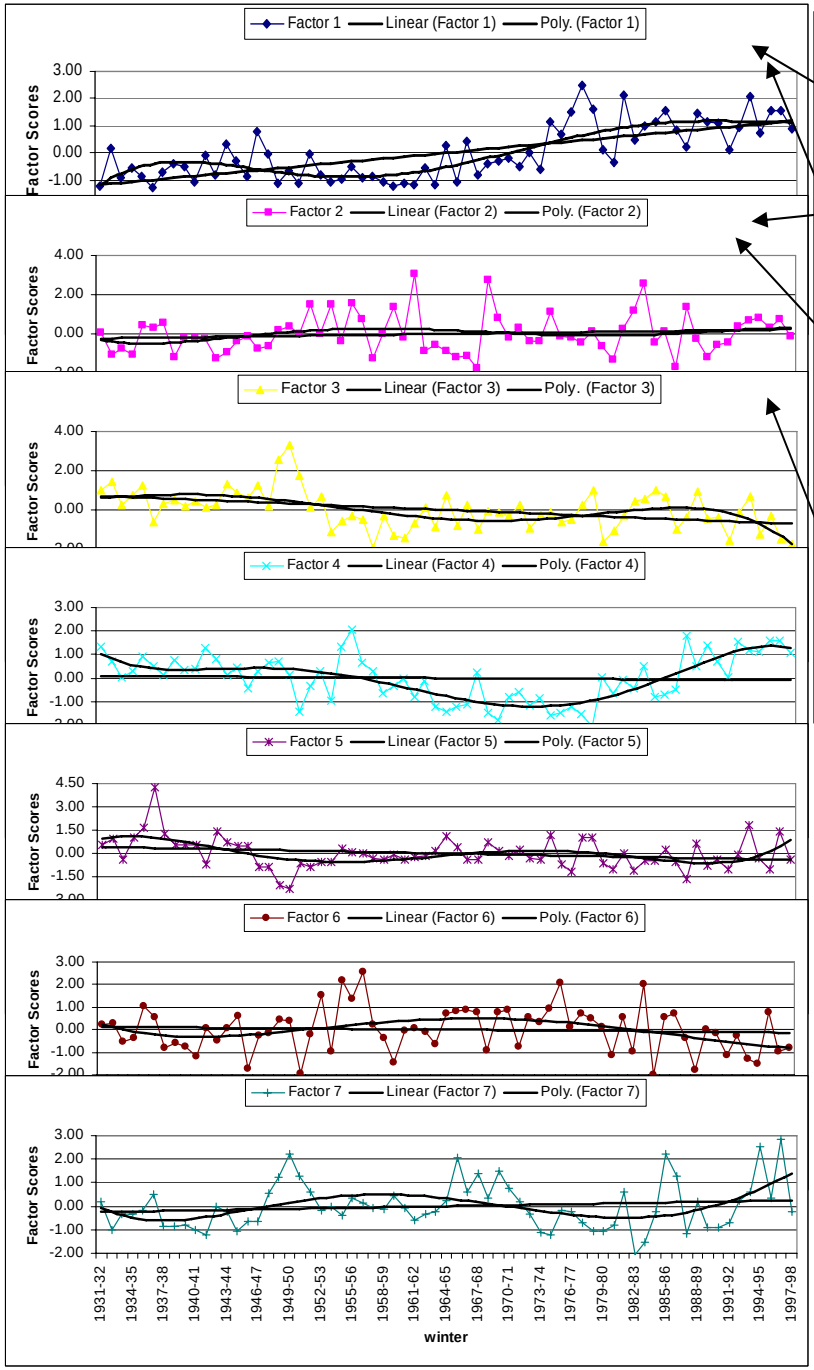
Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

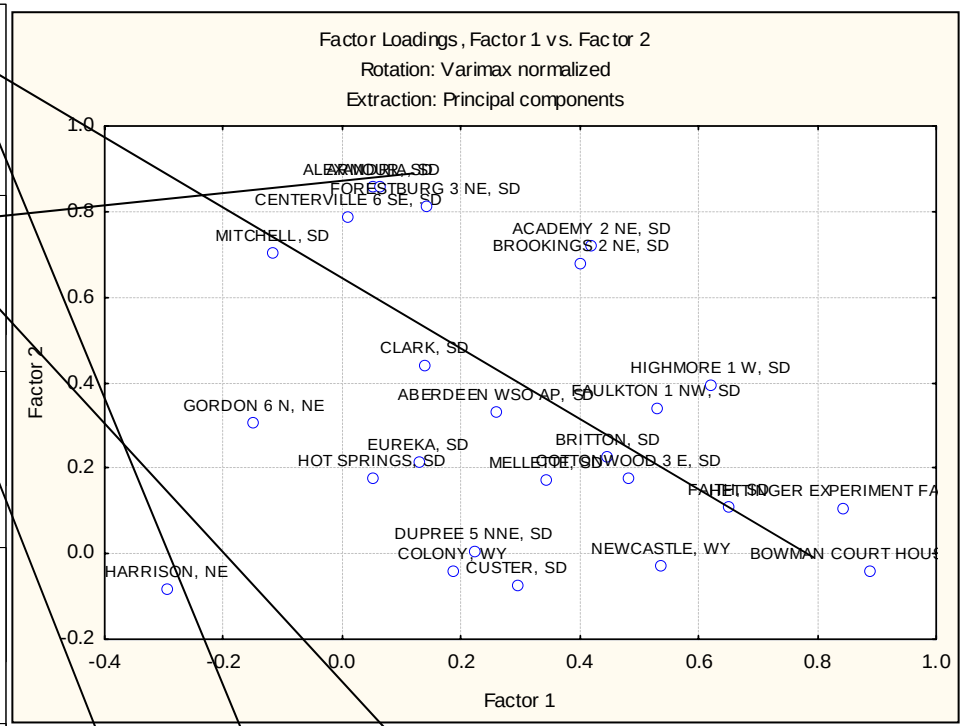
Regime of snow fall in SD

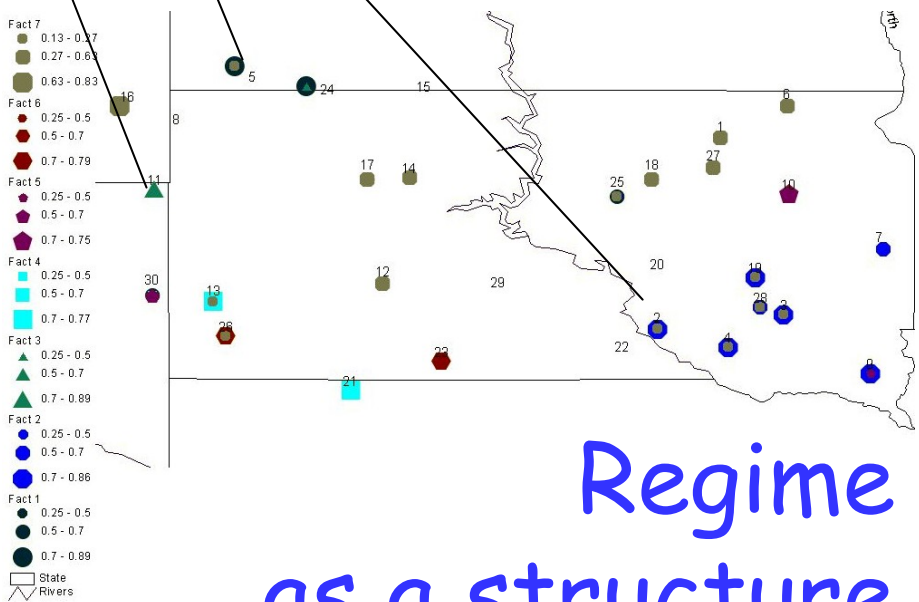


The initial matrix {Xt*n}, where are: t = number of years (67- 1931-1998) & n = number of meteorological stations (25),

Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008



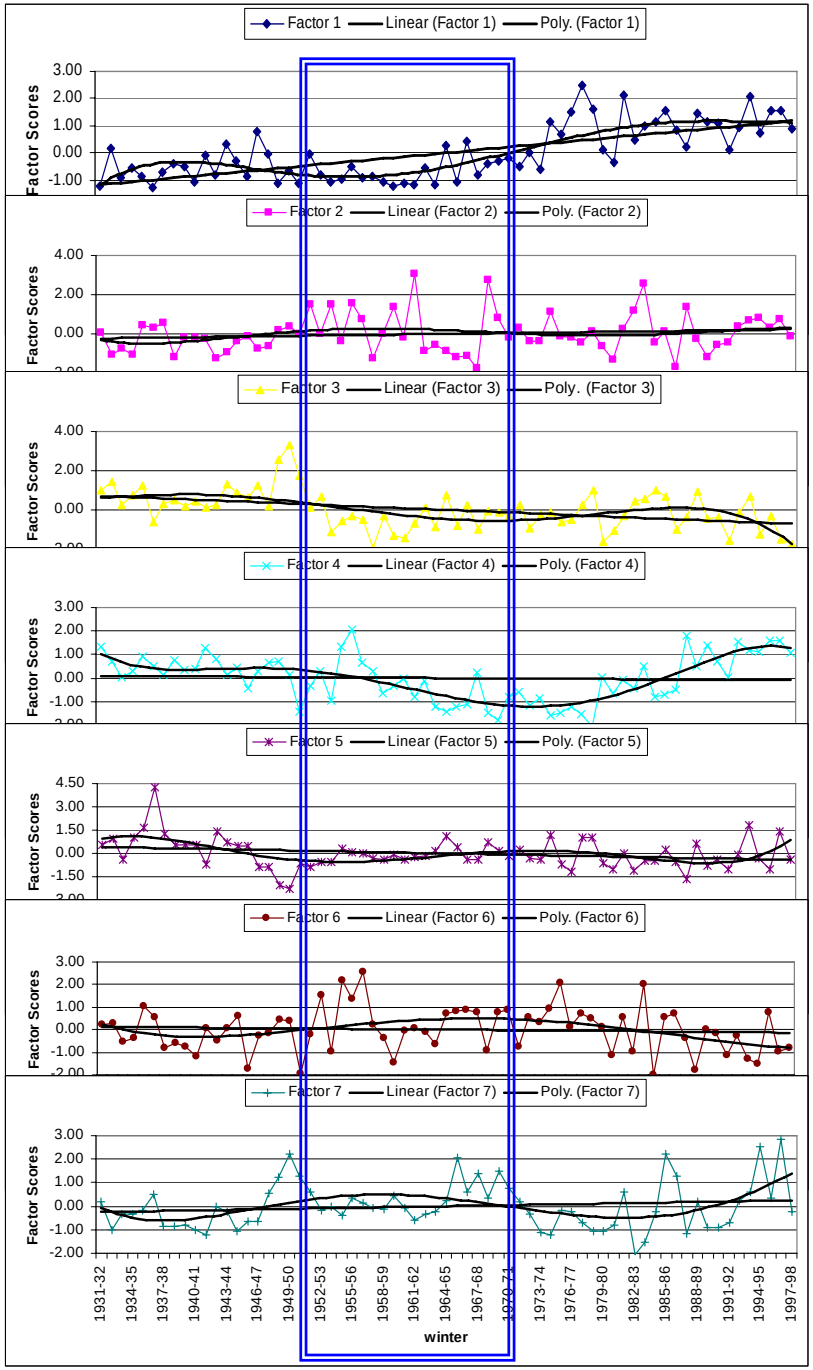


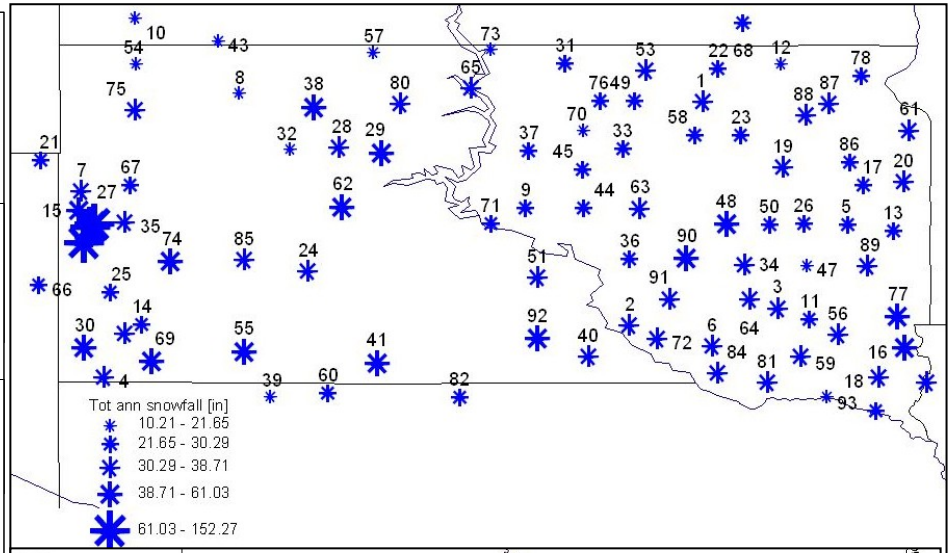


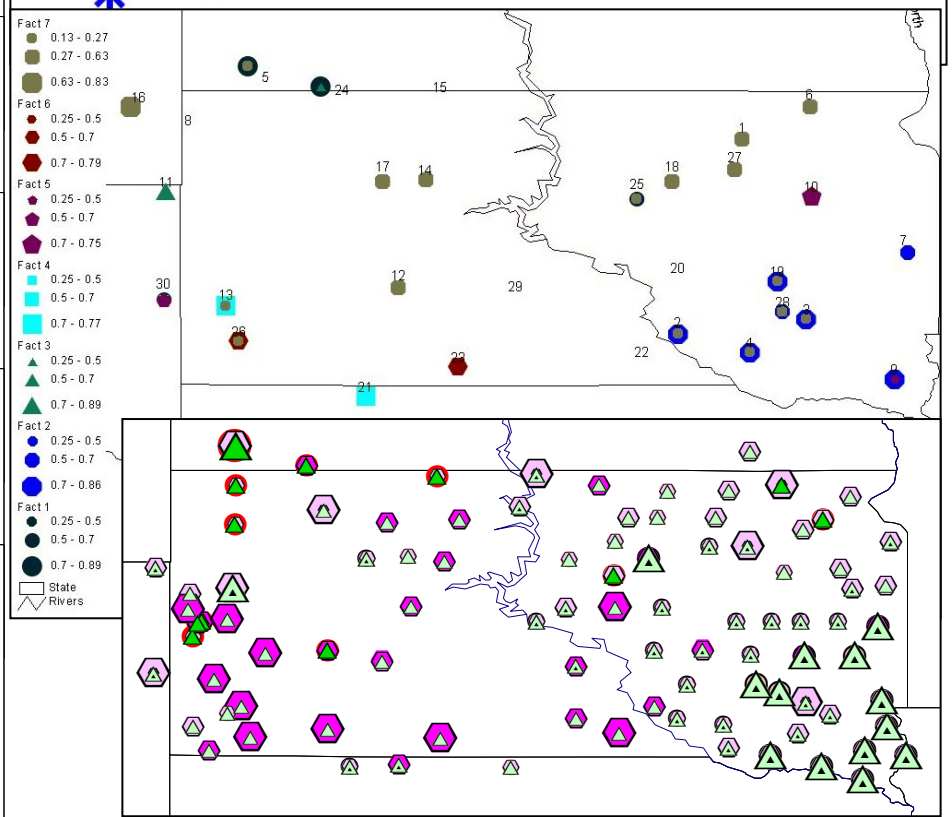
Regime as a structure

Regime & snow fall distribution

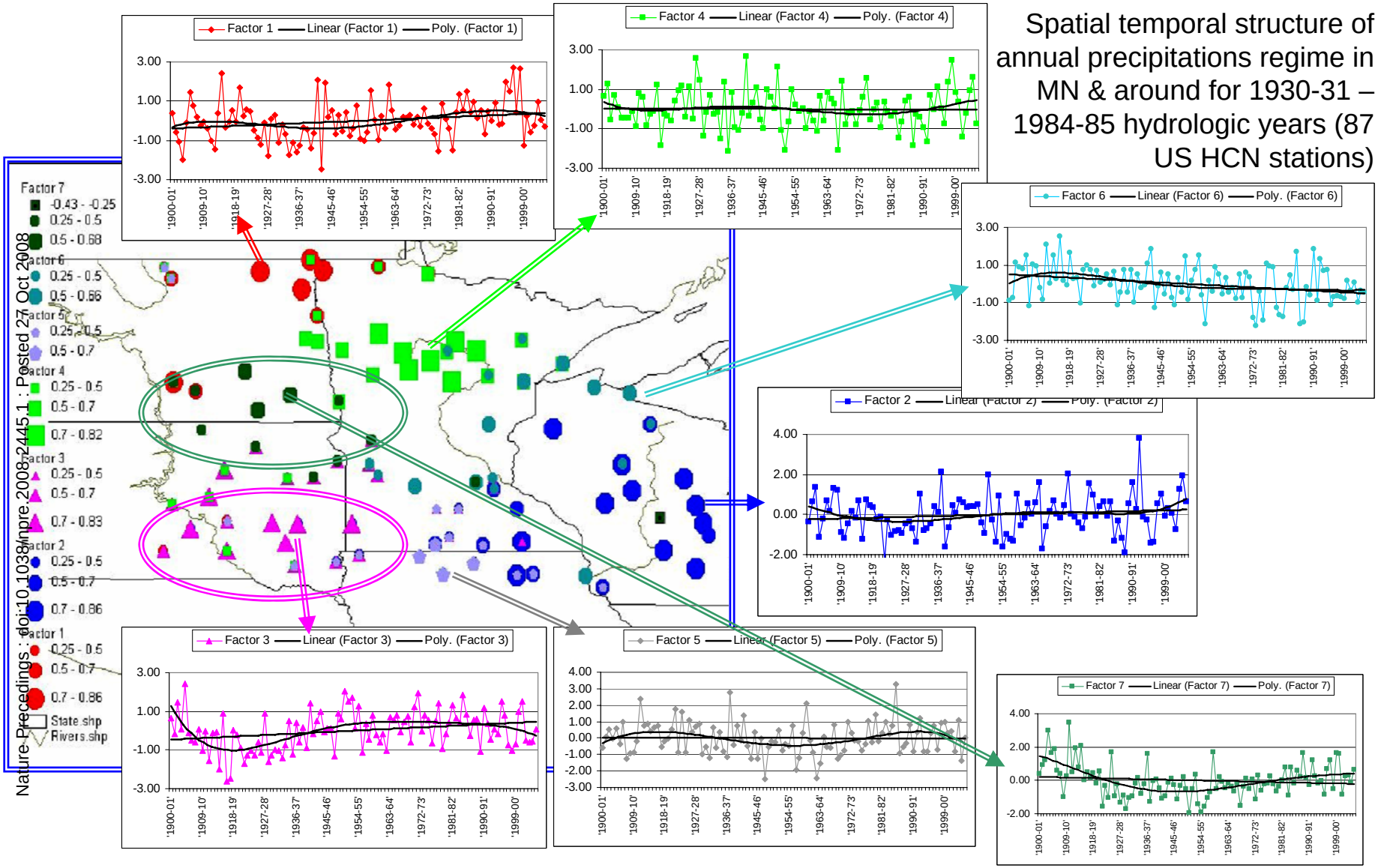
Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 22 Oct 2008







Spatial temporal structure of annual precipitations regime in MN & around for 1930-31 – 1984-85 hydrologic years (87 US HCN stations)



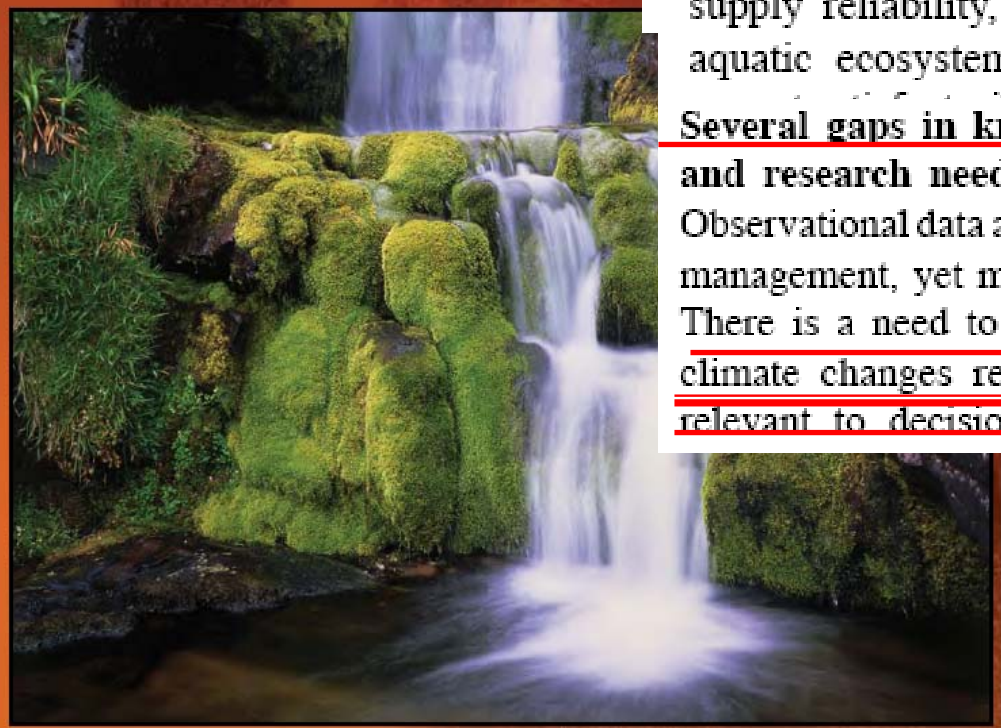
Nature Precedings : doi:10.1038/npre.2008.2445.1 : Posted 27 Oct 2008

Precipitation regime around SD

Improving the knowledge & creating maps of natural resources of SD

CLIMATE CHANGE AND WATER

IPCC Technical Pa



Intergovernmental Panel on Climate Change

WMO

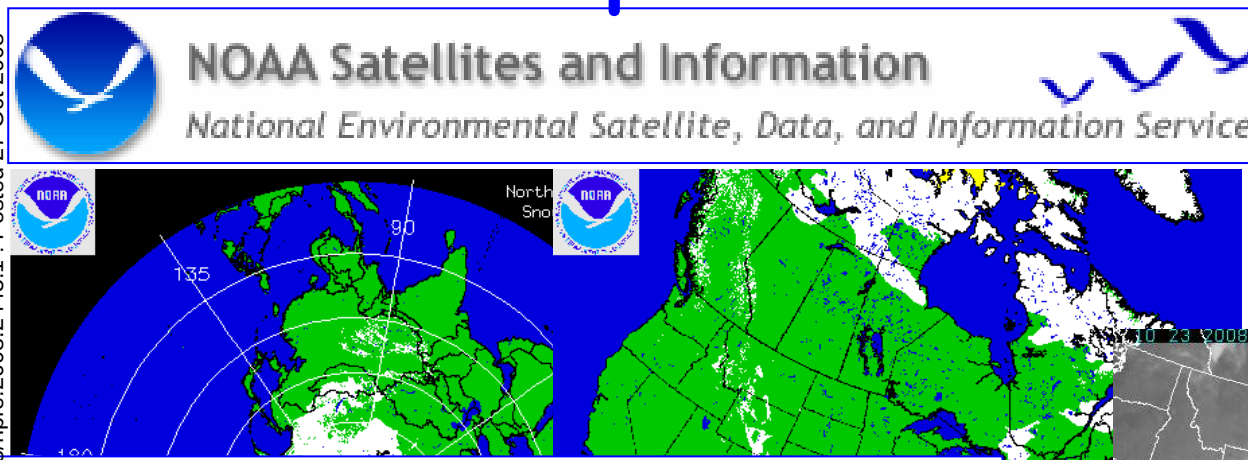
UNEP

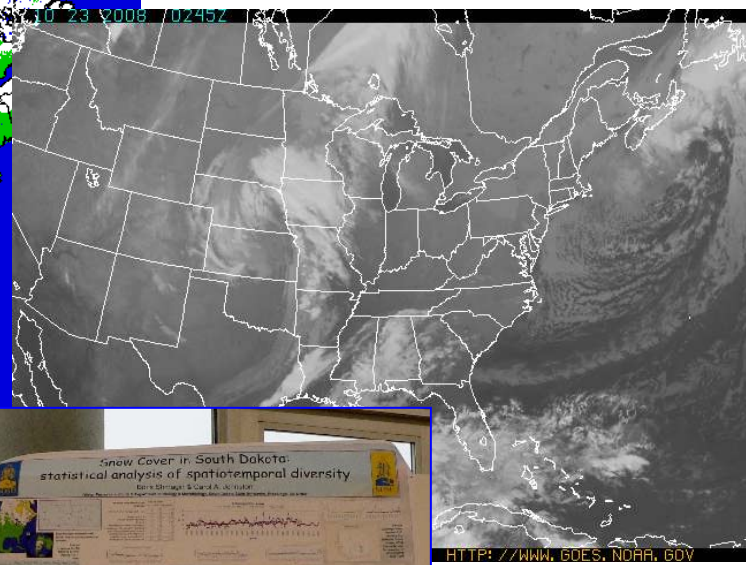
Observed warming over several decades has been linked to changes in the large-scale hydrological cycle such as: increasing atmospheric water vapour content; changing **Current water management practices may not be robust enough to cope with the impacts of climate change** on water supply reliability, flood risk, health, agriculture, energy and aquatic ecosystems. In many locations, water management

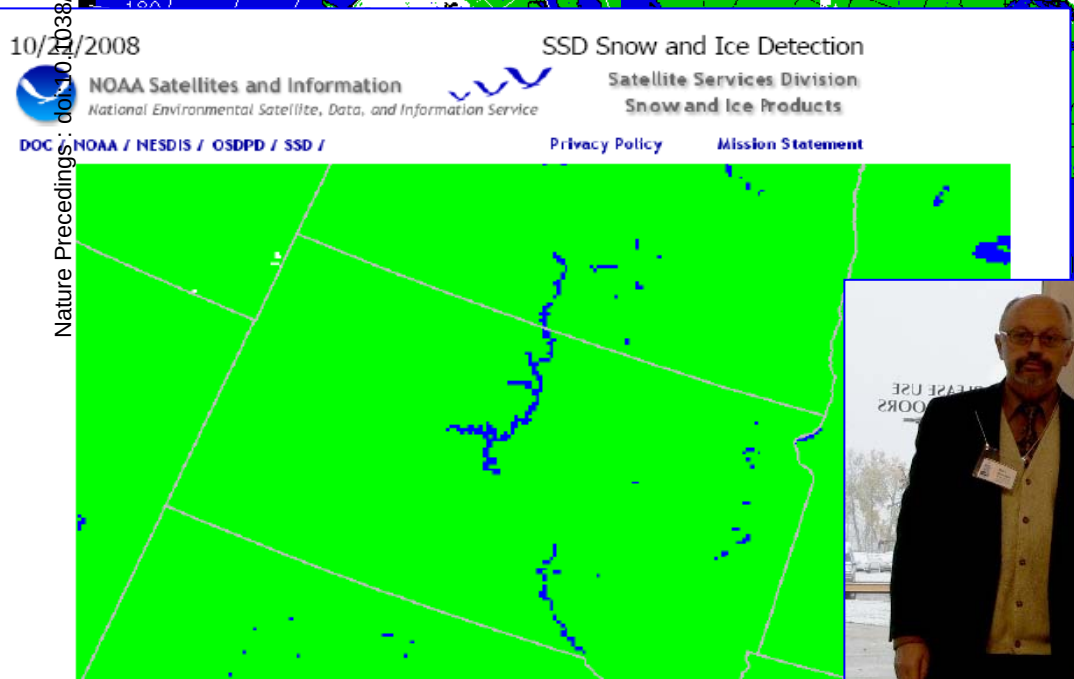
Several gaps in knowledge exist in terms of observations and research needs related to climate change and water. Observational data and data access are prerequisites for adaptive management, yet many observational networks are shrinking. There is a need to improve understanding and modelling of climate changes related to the hydrological cycle at scales relevant to decision making. Information about the water-

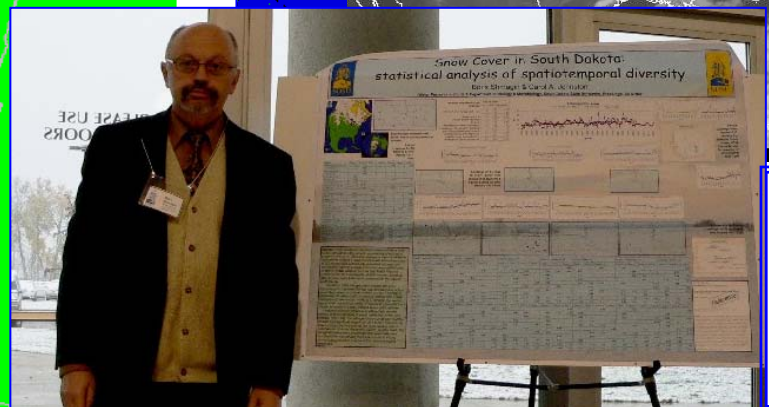
From Reports' conclusions

"… gaps in knowledge …" - from Reports' conclusions



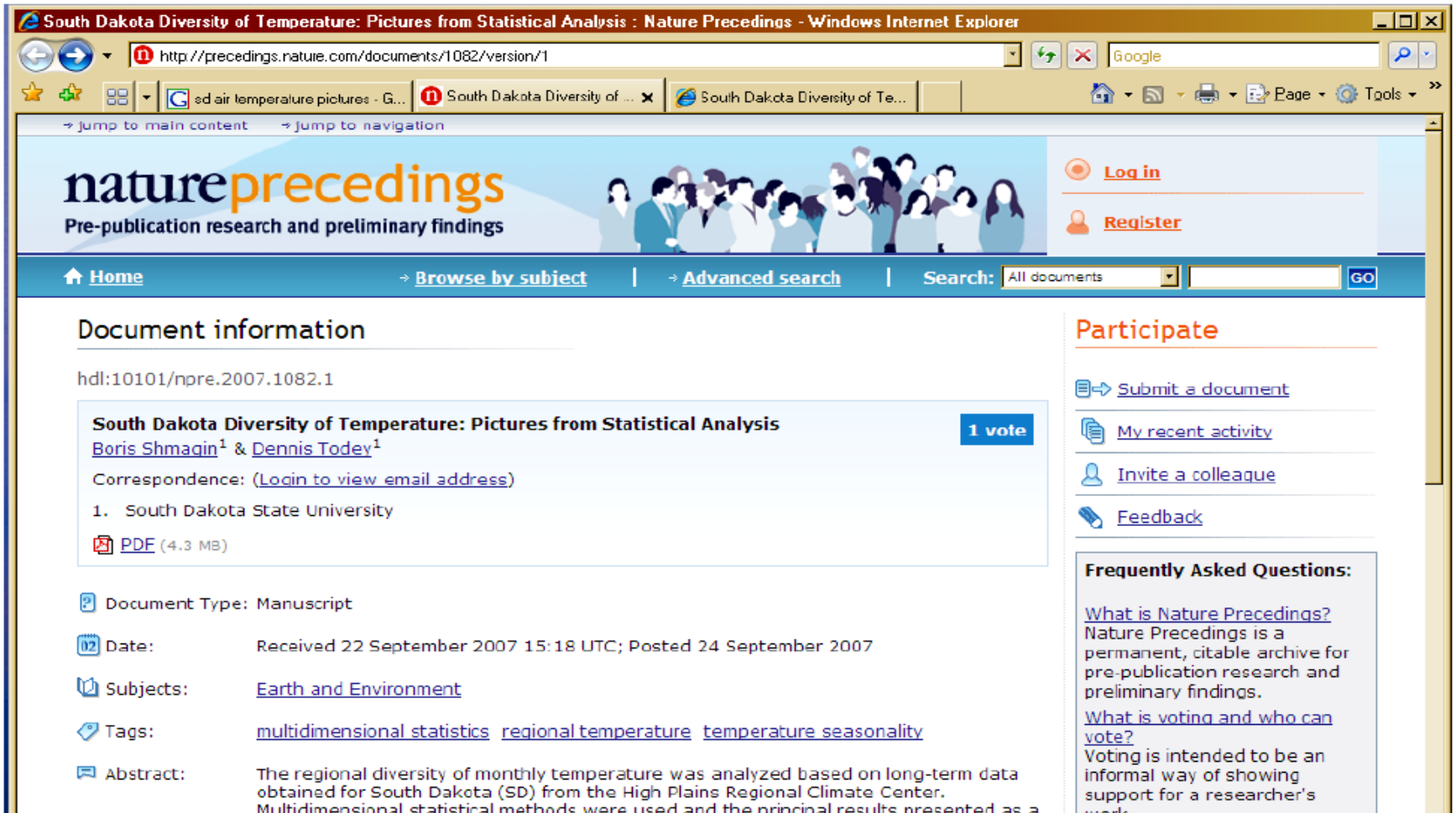






It was 1.5 inches of snowfall in Brookings, SD, on the day of poster presentation, Oct 22. 2008.

"... at scales relevant to decision making.." – from Reports' conclusions



Not from the Reports' conclusions

Academic Course as Research Seminar in Data Analysis on South Dakota Climate & Water Resources (for undergrad & graduate students)

Boris Shmagin & Din Chen

Department of Agricultural & Biosystems Engineering & Department of Mathematics and Statistics , South Dakota State University, Brookings, SD 57007

Abstract:

Data analysis allows quantitative description of the behavior of natural and human dynamic systems & opens a way to understand their structure. Seminar style academic courses are fundamental blocks for creating the craft atmosphere & research culture in the university. The originality of a realized approach is based on the use of special prepared reasonable small initial matrix ($X_{4\times7}$) with special constricted properties. This matrix is used to introduce all main statistical conceptions & techniques for students (mean, variability, correlation – as descriptive statistics, & also clustering and factorization – as multidimensional and multivariable exploratory techniques).

The second novelty are the illustrations based on the webpage (http://www.socr.ucla.edu/) to visualize the associated statistical principles of probability & distribution along with examples of real data that is a more appropriate way for students with non-mathematical majors grasp the concept presented. The designed course introduces models & tools for analysis of empirical climatic, hydrologic or environmental (economic, medical &d others) data for students & then provides the hands on experience involving techniques of data analysis for research.

The goal for students will be to produce a publication, oral or poster conference presentation or other presentation in accordance with the level of completed research (undergraduate, master or PhD). The student's reports will be the basis for evaluating success in learning. Analysis completed by students may be used in civil and ag engineering projects in SD for improving the normative base in the state, & also in research that require descriptions of SD's natural conditions like agriculture & wildlife.

SD Climate & Water Resources:
Analysis of Existing Empirical Data

I. PREFACE - - Course as a seminar for undergraduate & graduate students
II. INTRODUCTION – Empirical Data & Statistical analysis
III. PRIMARY STATISTIC ANALYSIS
1. Table of data (matrix of the initial model)
2. Parameters of sampling from five definitions.
3. Table of the parameters of variables.
4. Variable histogram.
5. Histograms of standardized data.
6. The graphs of variables constructed from the number.
7. Sampling nonuniformity definition according to Student's & Fisher's tests.
IV. PAIRED REGRESSION AND THE CORRELATION MATRIX
1. Simple linear model
2. Variance of the dependent variable and the coefficient of paired correlation.
3. A model of the dependent/independent variable ratio.
4. Regression & correlation for four pairs of variables.
5. Regression & correlation for the entire set of variables.
6. Correlation matrix.
7. Correlation matrix for samples from 15 definition.
V. MULTIPLE REGRESSION & CORRELATION
1. Multiple regression model.
2. A step-by-step regression for X4SR from 7 remaining variables.
3. The results of the step-by-step regression for the variables X4SR & X10R from the rest.
VI. CLUSTER ANALYSIS
1. The models & procedures of cluster analysis.
2. A hierarchical tree diagram of variables.
3. A hierarchical tree diagram of observation points.
VII. FACTOR ANALYSIS
1. A model of factor analysis.
1.1. *The method of component analysis.*
1.2. *The method of factor analysis.*
1.3. *Factor model parameters.*
2. The factor structure of the interrelationships between variables.
3. Distribution of sampling points by factor values.
3.1. *Graphs of factors and variables from the respecting groups.*
3.2. *Graphs of variable regression from factors.*
3.3. *Diagrams of point distribution in factor planes.*
VIII. CONCLUSION AND PRESENTATION OF EVERY STUDENT RESULTS ON SEMINAR

Professors:
Dr. Boris Shmagin redevelops course for SD on base of similar course; "System analysis of hydrological data" he was teaching for graduate students in hydrology for eleven years (1985-1996 Department of Hydrogeology Faculty of Geology Moscow State University, Russia. For course were developed & published textbook), & for one time (Ged 5281 Winter Quarter 1998 with Dr. Howard Mooers) in UMN-Duluth.
Dr. Dennis Todey is the state climatologist and regularly deals with practical issues of climatic data, including quality control & analysis of trends in such data. He teaches a 300 level class in climate & teamy-teaches a graduate level class in microclimatology.
Dr. Din Chen is Professor in Biostatistics in Department of Mathematics & Statistics, South Dakota State University

For discussion

Regional analysis of empirical hydro & climate data has to be the direction of SDSU research.

The goal is Atlas of climate & natural resources of South Dakota.

Acknowledgement
Research project is privately funded.

Additional support & funding for presented research provided by
Prof. Carol Peterson, SD EPSCoR &
Water Resource Institute of SDSU