Blended Learning for Disaster Risk Management in India – 2010 / 2011
Purpose of the Course

Overall Purpose

The purpose of the course in general and of the different modules in particular is to deepen and broaden existing knowledge of Statistical Terminology in the context of Disaster Risk Management, especially in India.

The online course - Disaster Risk Management

- Why do we engage in the use of Statistics for Disaster Risk Management?
- What can and - equally important – cannot be done with the use of Statistics for Disaster Risk Management?
- What are the basic concepts, terminology and tools for the use of Statistics for Disaster Risk Management?

The online course - Why collecting data?

We wanted to describe the reason and the methods to collect data and to improve knowledge and decision making based on information.

Collecting data is not without reason and the purpose here is to relate data collection and the use of information to the goals of Disaster Risk Management.

The online course - Why analyze data?

We wanted to describe the reason and the methods to analyze data and to improve knowledge and decision making based on these information.

The course is structured in three subject groups (Modules)

- Use and interpret Basic and descriptive Statistics
- Use statistical information on recent and passed disasters
- Understand “Cause and Effects” of disasters described by statistics

The online course - The learning process

- Find the appropriate data.source
- Understand and compare measures and techniques.
- Discuss outcomes and interpret results

The online course - The potential outcomes (1)

- To show that on data use and decision making there are potentially conflicting objectives and information
- Interpretation of statistical data can lead to stronger arguments for DRM, example: when it comes to seeking for additional funds for mitigation measures

The online course - The potential outcomes (2)

- Illustrate the complex nature of decisions making: e.g. the consequences of a flood are not only determined by the quantity of rainfall (meteorological aspect), but also by the capacity of the soil and the rivers to absorb the rain (topographical aspect) and the current type of land use (fields vs. settlements) and others
- Learn to identify the role of statistical data and analysis in the decision process and the need of an integrated program of facts recognition to improve decision making for DRM

The online course - Tools

- Self-paced interactive learning in Modules and units
- Modular (no linear progression, e.g. you can skip a module)
- Pretest and Self Assessments (Control your own progress)
- Tutored (Tutor(s) is/are available for explanation and requests)
- Chats (communicate with tutor and colleagues)
- Forum (Pin board or discussion board for exchange)
- WebQuest to apply acquired knowledge

The online course – Evaluate progress

Organizers and tutors can supervise progress of participants in .

- Time of participation in interactive learning in Modules
- Participation in Pretest and Self Assessments
- Participation in Chats
- Participation on Forum and discussion boards for exchange
- Delivered WebQuest to apply acquired knowledge

Thank You for your active participation in the course and your Attention!

Summary of Statistical Methods applied in the course and relevant for the WebQuest session

-

A Review

Content of this Presentation

Basics to understand Causes and Effects of Disasters as described by Statistics

Elements of Correlation and simple linear Regression

Examples of multiple linear Regression

What is the purpose?

Summary of Statistical Methods applied in the course and relevant for an interactive “WebQuest” session:

Topic of the Exercise: (1) How to make valuation of loss due to excessive rain in a district and (2) data sources to be used?

Refer to M3, WebQuest 1 : Understand Causes and Effects of Disasters as described by Statistics?

Something done online - but not reviewed

Almost all referred text is available online

Descriptive Statistics for Samples

- Discrete Example
- Continuous Example
- Centre of a Distribution
- Comparison of Mean, Median, and Mode
- Spread of a Distribution

Some Basics: Sampling

We may restate the definition of simple random sampling in more mathematical terms for future reference:

A simple random sample is a sample whose n observations X1, X2, . . . , Xn are independent. The distribution of each Xi is the population distribution (with mean μ and variance σ2).

The Central Limit Theorem

The central limit theorem: As the sample size n increases, the distribution of the mean $\bar{x}$ of a random sample taken from practically any population approaches a normal distribution (with mean μ, and standard deviation)

$\mu = \bar{x} \pm$ a sampling error

The crucial question is: How wide must this allowance for sampling error be?

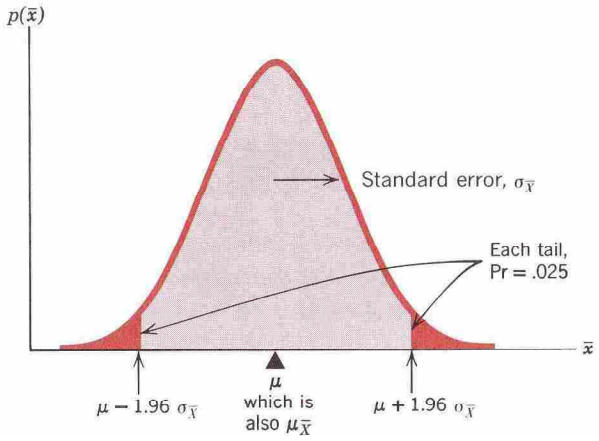
Sampling Errors

We may restate the definition of simple random sampling in more mathematical terms for future reference:

A simple random sample is a sample whose n observations X1, X2, . . . , Xn are independent. The distribution of each Xi is the population distribution (with mean μ and variance σ2).

The Sample and the Bell curve

Normal distribution of the sample mean around the fixed but unknown parameter μ. 95% of the probability is contained within 1.96 standard errors



Necessary Characteristics of a Sample

A Sample has to be

- Sufficiently big (enough sampled elements)
- It has to contain randomly selected elements

After this we can make tests (Ho = no difference) or predictions about mean and variance

This does not tell anything yet about a validity of a prediction

Cause and Inference

Cause / Causality is the relationship between an event (the cause) and a second event (the effect), where the second event is a consequence of the first. Inference deals with associations, relationships, correlations, etc.--but NOT with causal connections--between variables. Causality cannot be measured, determined, or established by observation alone. Additional inputs are required to infer a causal connection or inference. So an associative inference might be causal or not

Our assumption about inference

We might determine (statistically) an inference between events (e.g. rainfall and losses) but this associative inference might be causal or not

Rainfall River Gauges Protective measures	Might be related but might not be caused. So we skip the word: Cause →	Losses Human Lives, Agriculture, Infrastructure, Economic

An assumption to prove statistically

There is an inference between

Rainfall River Gauges Protective measures	and	Losses Human Lives, Agriculture, Infrastructure, Economic

What to do?

1. Does the assumption make sense?
2. Collect data with respect to sample theory
3. Choose a statistical method / technique to discern inferences
4. Develop model
5. Can you verify the model in reality?

An exercise to prove this assumption

Does the assumption make sense?	✓
Collect data with respect to sample theory See Rainfall&LossesBihar.xls from M3U2	(30 years of rainfall and casualties) ✓
Choose a statistical method / technique to discern inferences	**Correlation**
Develop model	**Linear Regression**
Can you verify the model in reality?	To be shown

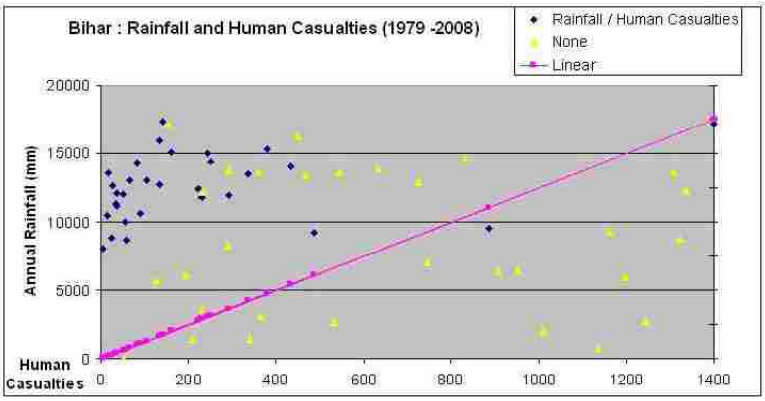
Correlation (1)

Correlation is used frequently in the very general sense that two or more variable depends from each other somehow.

Let's assume we have an indicator of the magnitude of disasters (e.g. number of casualties) and another variable, e.g. the rainfall again. Let us try to put these two variables in a scatter plot

Correlation (2)

The Scatterplot



Correlation (3)

The Pearson's correlation coefficient r

Correlation is a measure of the relationship between two numerical Y variables, or sets of variables.

The correlation coefficient (r) is a measure of the strength of the correlation; it varies from -1 (perfect inverse correlation) through 0 (no correlation) to +1 (perfect positive correlation).

Correlation (4)

The Results of the correlation matrix

Correlation Matrix (R)		
	Human Casualties	Annual Rainfall (mm)
Human Casualties	1.000	0.296
Annual Rainfall (mm)	0.296	1.000

So we conclude from plot and calculation, that there is a weak relationship between the two variables but far from the assumption that rainfall alone is “causing” floods in Bihar. (A correlation greater than 0.8 would be described as strong, whereas a correlation less than 0.5 would be described as weak

Linear Regression (1)

The linear regression explains the variation in one variable, a dependent numerical variable (Y), in terms of one or more predictor variables, independent numerical variables (X1, X2, ... , Xm). It assumes a linear (straight line) relationship between the dependent variable and the independent variable(s) of the form

$$Y = a + bX$$

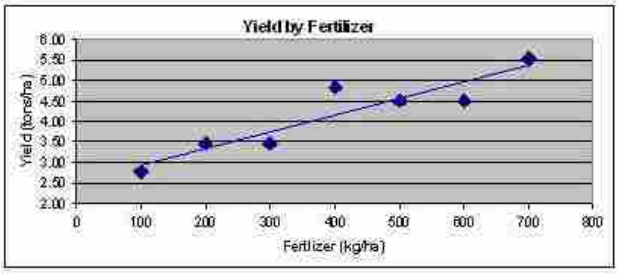
Linear Regression (2)

As a simple mathematical model, it will be useful as a description, or as a means of predicting the yield Y for a given amount of fertilizer X. Initially we restrict the discussion exclusively to how a straight line may best be fitted.

Since yield depends on fertilizer, yield is called the "dependent variable" or "response variable" Y. We refer to fertilizer as an "independent variable" or "factor," or "regressor" X .

Linear Regression (3)

X Fertilizer (kg/ha)	Y Yield (tons/ha)
100	2.748
200	3.460
300	3.460
400	4.844
500	4.498
600	4.498
700	5.536



the general Formula of the trend or regression line will be $\hat{Y} = a + b*X$, a is called the intercept, b the slope

For the example $\hat{Y} = 2.521 + 0.0041*X$ is calculated

Linear Regression (4)

If our model is a good explanation of the relation between the independent and dependant variables, you can use the coefficients to forecast the value of the dependent variable. (b = increase in Y if x is increased by one unit)

The Coefficient of determination (R2) explains by the variability in the dependent variable(s) X. Usually this is an indicator of the “quality” of the model. The adjusted R2 should be used for comparisons of different regressions.

Linear Regression (5)

Multiple Regression

Multiple regression is the extension of simple regression, to take account of more than one independent variable X.

In our example a better prediction of yield may be possible if both fertilizer and rainfall are examined.

Usually the adjusted R2 again is an indicator of the “quality” of the model.

Linear Regression (6)

An example of Multiple Regression:

Regression based on the model whereby Rainfall and Administrative measures supposedly explain Drought Disasters in Gujarat:

Solutions_DroughtResultsGujarat1982-2008Ex.xls

Year	No. of Districts Affected	No. of Villages Affected	MAM	Minor Irrigation Schemes	Soil & Water Conservation Schemes	Area Development Programme
1998	24	4633	0	134.3	46.5	21.3
1999	20	20069	217	170.3	61.2	27.2
2000	26	23406	567	212.8	79.6	23.4
2001	31	30583	262	195.9	88	16.6
2002	18	7964	5	102	58	6.5

DataForRegression (6) | Regr.-RFs | Regr.IrrigationYields | RF+Yields2 | RF+Yields2a

Linear Regression (7)

This exercise led us to data including projects on improving the water resources of the state and there we obtained satisfying results (independent exercise 2):

Summary			
	R^2	R	Adj. R^2
	1.000	1.000	1.000

A good fit of the model

A common sense explanation: Pre monsoon rainfall and improving the water resources of the state influence the drought disaster results in Gujarat

Linear Regression (8)

"Partial Regression Coefficients indicate how much the dependent variable will change (per unit) if the corresponding independent variable is increased, if all other independent variables are held constant".

Regression Coefficients

Source	Coefficient	Std
Intercept	-43655.345	
MAM	-3.118	
Minor Irrigation Schemes	-322.464	
Soil & Water Conservation Schemes	1291.368	
Area Development Programme	1481.063	

Linear Regression (9)

Regression Coefficients

Source	Coefficient	Std
Intercept	-43655.345	
MAM	-3.118	
Minor Irrigation Schemes	-322.464	
Soil & Water Conservation Schemes	1291.368	
Area Development Programme	1481.063	

This would mean: The unit of the dependent variable is one village.

- If we have a decrease of 1 mm of pre monsoon rainfall we can expect 3 more village to be affected by drought.
- If we decrease our Minor Irrigation Schemes by 1crore Rupees per year we can expect 322 more village to be affected by drought. Not necessarily the contrary will hold true increasing the output to reduce the number of affected villages but it is very likely

Linear Regression (10)

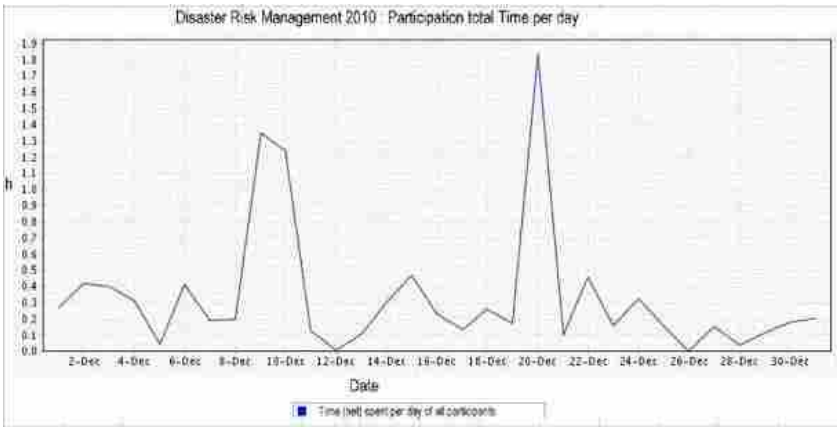
There are obviously some limits to this model (small data range, explanation of the remaining variables) but the use of fact based information to explain causes of droughts in Gujarat might be more transparent to you than before.

Please bear in mind: We have constructed these models purely to show you the statistical tools. Once you are familiar you will have the opportunity to apply them in your professional field and you can refine the models, the underlying concepts and improve with the help of your own organization the quality and consistency of your data, thereby improving measurements and finally the analysis.

Thank You for your active participation in the course and your Attention!

Blended Learning for Disaster Risk Management in India – 2010 / 2011
Assessment of the Online Course

Overall Participation December 2010



Overall Participation December 2010 Comments

The overall participations show high peaks of participation for the weekends or on Christmas holidays and/or vacation which indicates time constraints of participants due to other obligations.

Participation Statistics and Progress

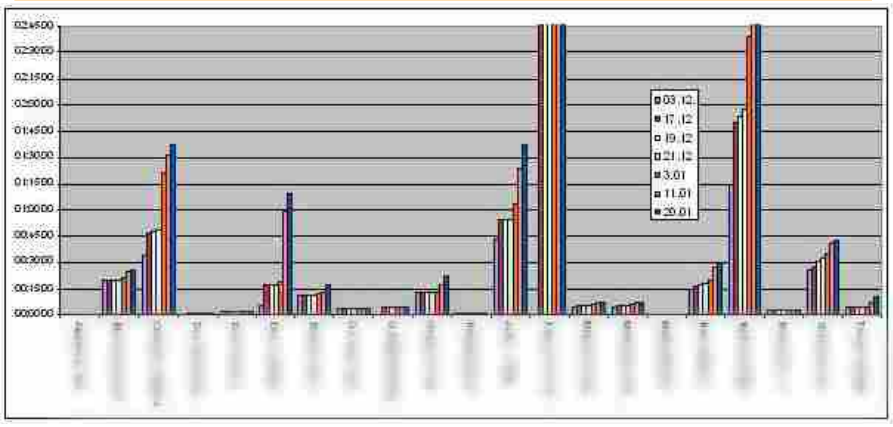
Participant	Active Time by 19.12.2010 hh:mm:ss		Active Time by 20.01.2011 hh:mm:ss	
Aggarwal, Shri Mayank (mayank-aggarwal)	00:00:00		00:00:00	
[redacted]	00:21:25	✓	00:25:33	✓
[redacted]	01:21:15	✓	01:37:50	✓
[redacted]	00:00:58		00:00:58	
[redacted]	00:01:22		00:01:22	
[redacted]	00:18:59		01:09:13	✓
[redacted]	00:11:37		00:16:46	
[redacted]	00:03:47		00:03:47	
[redacted]	00:03:55		00:03:55	
[redacted]	00:13:15		00:22:03	✓
[redacted]	00:00:57		00:00:57	
[redacted]	01:03:49	✓	01:37:57	✓
[redacted]	05:19:35	✓	05:45:57	✓
[redacted]	00:05:56		00:06:38	
[redacted]	00:05:56		00:06:38	
[redacted]	00:00:34		00:00:34	
[redacted]	00:19:38		00:29:31	✓
[redacted]	02:39:12	✓	03:06:03	✓
[redacted]	00:02:46		00:02:46	
[redacted]	00:34:33	✓	00:42:51	✓
Talwar, Shri Madhwarappa (madhwarappa-talwar)	00:03:51		00:10:36	
Number of Parrticipants active> 20 min		**6/21**		**9/21**

Participation Statistics and Progress Comments

The participation of the online course was very heterogeneous. We count 21 participants.

The time participating ranges from flat 0 to more than 5 hours and these times are net time participation. All participants with a net time over 20 Min (indicating some constant activity) have been given a tick mark.

Progress of Participation by Participant



Progress of Participation by Participant Comments

The graph is showing participation at various stages of the course one could notice a slow start in many cases, some participants even started late in December. And what is also surprising a constant growing activity until the end (10.01.2011) and even beyond as the increasing magenta and blue bars show in the bar graph. This remark is clearly limited to the mentioned 9 participants out of 21

Pretest and Self Assessments

Pretest	First number: finished and submitted, second number: begun but not submitted
As of 03.01.2011	As of 20.01.2011
Finished and Submitted (11+3/22)	**Finished and Submitted (15+1/22)**

Self Assessment 1	
As of 03.01.2011	As of 20.01.2011
Finished and Submitted (5+1/22)	**Finished and Submitted (10+1/22)**

Self Assessment 2	
As of 03.01.2011	As of 20.01.2011
Finished and Submitted (4+1/22)	**Finished and Submitted (8+0/22)**

Self Assessment 3	
As of 03.01.2011	As of 20.01.2011
Finished and Submitted (3+0/22)	**Finished and Submitted (5+0/22)**

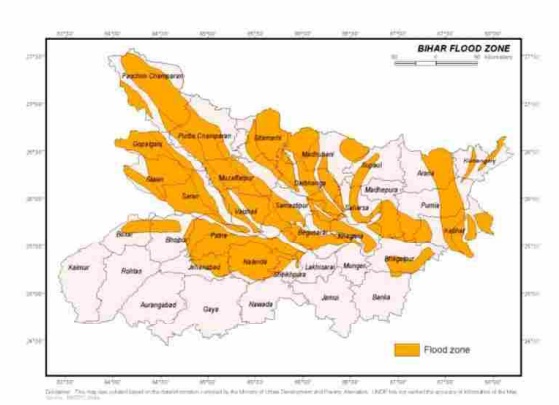
Pretest and Self Assessments Comments

For the Tests and Assessments there is a comparative picture, the Pretest has been delivered by 15/22. This was not so much of a test but an initial appetizer for the interactive character of the course. It shows a lot of activity and judging by the results All 15 respondents have basic or advanced knowledge of statistics, all it needs to enter and do the course with realistic expectation of benefit.

Pretest and Self Assessments More Comments

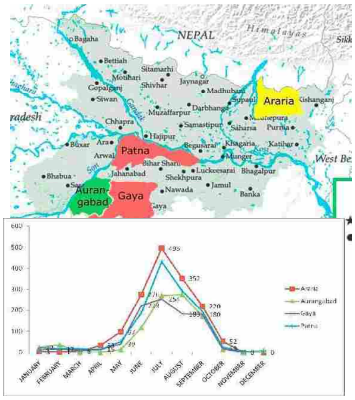
The number of deliveries of Self-Assessments constantly decreases with increasing difficulty of the modules but a total of 5 participants have continued until the self assessment of the 3rd Module. The self assessments are self triggered and they do not allow to evaluate the learning achievements. However and again participation has increased during the last weeks and a substantial improvement of number of deliveries is noticeable even during the last two weeks.

Delivered WebQuests (1)



A scheme for the districts of a state (e.g. Bihar) to be classified as low, medium or high risk district regarding rainfall related disasters
Sreeja Nair

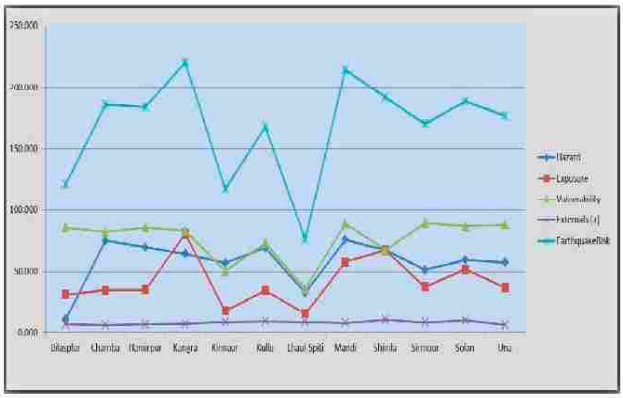
Delivered WebQuests (2)



Identify the seasons and months of the year, when prevention and preparedness actions will be necessary for the chosen districts in Bihar

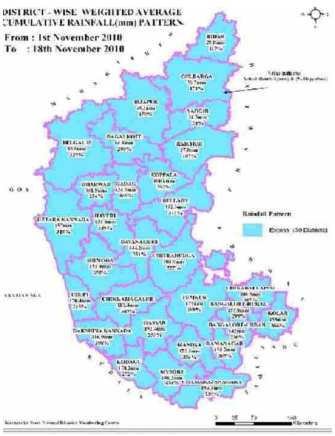
Identify the differences among the districts in rainfall patterns, peaks, length and distribution and Look for arguments to prove, that flood patterns have changed over the years in Bihar :
Hemprabha Chauhan, Prithwi Jyoti Bhowmik, and Dr. Pradip Kr. Nath

Delivered WebQuests (3)



Status of Earthquake Vulnerability and Disaster Risk Management In Himachal Pradesh
Dr. G.P. Kapoor,

Delivered WebQuests (4)



Rainfall Pattern And Disaster Risks in Districts of Karnataka State
Mr.Talmar

Delivered WebQuests
Comments

Six participants have delivered the WebQuests by the 24.01.2011. A WebQuest should expose the experiences of the participant (or the group of participants) in the area of Disaster Risk Management blending these experiences with the knowledge transmitted during the course. The 4 delivered WebQuests, one was elaborated by a group of 3, have all performed accordingly and the goal of the WebQuest has been reached by all.

Delivered WebQuests
More Comments

The first two delivered WebQuests apply both the techniques of observing and analyzing meteorological observations for the forecasting and mapping of flood risks in Bihar. These solutions stick close to the techniques exposed in the first module of the online course and add an innovative approach to the presentation. Whereas the first bemoans the insufficiency of met. data for flood prediction ventures the second into trying new and practical solutions.

Delivered WebQuests and More Comments

Another WebQuest was chosen to assess the hazard risks for the Karnataka State. This was an excellent presentation but missing some relation to the course content

One last solution of the WebQuest stands out, not only choosing the rather most demanding of the statistical methods PCA and Factor analysis but applying it to data sources and presenting the analytical results of a particular new area of work: Himachal Pradesh.

Conclusion

the course demanded some dedication and depending on the prior background knowledge of the participants some digging in new and unfamiliar soil. Several participants conquered new ground with amazing astuteness and perseverance. That they have done so, should fill the organizers, the administrators, the course designers and the active participants with great satisfaction.

Thank You for your active participation in the course and your Attention!

Agenda of the Face to Face Course

Purpose of the course:

1. Discussion of the e-learning modules, answering questions of the participants and getting feedback from the participants for revision of the course material
2. Presentation and discussion of the WebQuests elaborated by the participants
3. Practical application of the approaches and methods taught in the course (this requires additional preparation of the participants in advance to the workshop)
4. Planning of new blended learning courses to be developed in the future

Day 1: Thursday, January, 27th, 2011

Time	Subject	Protagonists
09:00 – 09:45	Welcome and introduction of participants	NIDM, Ifanos, GIZ; Participants
09:45 – 10:00	Opening address by coordinator of ex InWEnt / now GIZ	GIZ (Mr. Björn Kalscheuer)
10:00 – 10:30	Recapitulation of Goals of the course	NIDM, Klaus
10:30 – 12:00	Summary of Statistical Methods applied in the course and relevant for the WebQuest session	Klaus
12:00 – 13:00 14:00 – 15:30	Presentation of the delivered WebQuests	Authors
15:30 – 17:00	Discussion of the WebQuests	All

Day2: Friday, January, 28th, 2011

Time	Subject	Protagonists
09:00 - 09:30	Practical Application: Introduction of the topic Presentation of results of preparatory steps done by the participants (see below)	Klaus, Participants
09:30 – 13:00	Joint elaboration of the exercise (see below)	Klaus, Participants
14:00 – 15:00	Evaluation and Discussion of the current course	NIDM, Ifanos, GIZ, Klaus, Particpants
15:00 – 16:00	Planning new blended learning courses 1, required human and other resources and infrastructure, tools etc.	NIDM, Ifanos, GIZ, Klaus, Participants
16:00 – 17:00	Planning new blended learning courses 2, Proposals of forthcoming Blended learning courses	NIDM, Ifanos, GIZ, Klaus, Participants
17:00	Closing of the Workshop	NIDM, Ifanos, GIZ

Outline of the Practical Application Task

One of the main purposes of the face-to-face course is to apply the approaches and methods taught during the e-learning phase.
In order to be able to focus on the application part, the participants are asked to do some preparatory steps in advance to the workshop.

The following scheme introduces the topic to be worked on, the preparatory steps and the application steps which will be elaborated during the workshop.

Topic of the Exercise	**Can you suggest how to make valuation of loss due to excessive rain in a district and data sources to be used?**

Preparatory steps to be done in advance by the participants

The participants can use WebQuest 1 from Module 3 as an orientation for these steps (UNDERSTAND CAUSES AND EFFECTS OF DISASTERS AS DESCRIBED BY STATISTICS (http://gc21.inwent.org/ibt/login/GC21/area=module/main/en/wbt/gc21/disasterRiskManagement/module03/DRM_WebQuest_Mod3_1.pdf)

Step 1	Identify a region in India (state, district, etc.), where rainfall data and economic data are available for a minimum period of 5 years Result: Specification of a region and list of data sources
Step 2	Specify your questions based on the topic given above. What do you want to find out? What can you find out depending from the data available? E.g. information on which types of losses (e.g. agricultural, housing, indemnities paid) is available Result: List of specific questions which can be answered using the available data; e.g. How are losses correlated to rainfall or other meteorological data like river gauges? Or how are damages of roads or infrastructures correlated to rainfall and how could this be measured?
Step 3	Assess the available data sources and decide which data will be used for the analysis. Justify reasons of choice. Download the relevant data and bring it to the workshop Result: Data for application are available

Steps to be executed during the workshop

Step 4	Presentation of the results of Steps 1 to 3 by the participants
Step 5	Joint application of the techniques introduced in Module 3 (e.g. Correlation, Regression) to relate rainfall data to economic data (gains and losses).
Step 6	Presentation and discussion of results
Step 7	Valedictory session

Infrastructure required:
Internet access for the participants is required.

Saturday, January, 29th, 2011
Follow-up meeting between NIDM, GIZ, ifanos (approx. 2hrs)