

FOT-Net seminar on data management

Frank Lai

Institute for Transport Studies

University of Leeds

So why am I here?

To share some experience in data collection through field trials

- A fully instrumented car
- 20 ISA cars
- And the forthcoming 550 ISA cars

Overview of the talk

Raw data \Rightarrow database \Rightarrow statistical packages

- Data acquisition
- Data management
- Getting the data ready for analysis

Physical access to the car

- suits smaller scale trials
- ideal for large data files (i.e. video)
- ideal if regular access to the car is required
- very important – store the data on a file server!!

Wireless transmission

- suits larger scale trials
- ideal for smaller data files (i.e. text format)

Minimise the risk of data loss

- Monitoring of data collection
 - request the participants to report faults
 - secondary on-board storage
 - automated message sent to base upon system faults
 - automated summary reports at the end of a trip
 - attend to problems promptly

Minimise the risk of data loss

- Establish a rigid procedure for data handling

sub_id	veh no	data collection starts	data collection ends	Archived / Uploaded			
1	7	1230 25/03/04	1500 23/04/04	Sub01_PH1_Car07	KEY:		
	7	1500 23/04/04	1500 28/05/04	Sub01_PH2_M1_Car07			PH1
	7	1500 28/05/04	0730 25/08/04	Sub01_PH2_M234(1)_Car07			PH2
	7	0730 25/08/04	1400 26/08/04				PH3
	1	1400 26/08/04	1330 8/09/04		1600 12/8/03	1430 21/8/03	No fuses i
	20	1330 8/09/04	1450 8/09/04	Sub01_PH2_M234(2)_Car20	1430 23/9/03	1200 14/10/03	redundant
	20	1450 8/09/04	1515 29/09/04	Sub01_PH3_Car20	All time stamps are GMT		
2	21	1400 11/8/03	1600 12/8/03		Sub02_PH1_Car15		Data down
	21	1600 12/8/03	1430 21/8/03		Sub02_PH1_Car15		Data uploa
	15	1430 21/8/03	1500 23/9/03	Sub02_PH1_Car15	Sub02_PH1_Car15		Data uploa
	15	1500 23/9/03	1630 8/10/03				
	12	1630 8/10/03	1200 14/10/03				
	12	1200 14/10/03	1400 12/2/04				
	15	1400 12/2/04	1400 17/02/04				
	20	1400 17/02/04	1700 13/03/04	Sub02_PH2_M1(1)_Car20			

Minimise the risk of data loss

- Backup, backup, and backup

Also pay attention to data confidentiality

- Password protected files
- restricted access to the file server

Get the data ready for analysis

- Database structure
 - vehicle data (high resolution data, trip data, ...)
 - map data (road attributes, geographic data, ...)
 - driver characteristics (gender, attitudinal data,...)
 - etc.
- Data reduction
 - data extraction
 - aggregate data according to analysis requirements

Thank you for your attention