Using Instrumented Probe Bicycles to Develop Bicycle Safety and Comfort Prediction Models

A. Lee, L. Dias, S. Mohanty, T. Carvalho, J. Commandeur, and G. Lovegrove

2014 International Cycling Safety Conference
Gothenburg, Sweden

a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA
Sustainable Transport Safety Lab

FACULTY OF APPLIED SCIENCE
SCHOOL OF ENGINEERING

SWOV INSTITUTE FOR ROAD SAFETY RESEARCH

SUSTAINABLE TRANSPORT SAFETY

URBAN systems

Dr. Gord Lovegrove

Sustainable-Transport Safety-Researcher

Associate Professor, School of Engineering

University of British Columbia, Kelowna, Canada

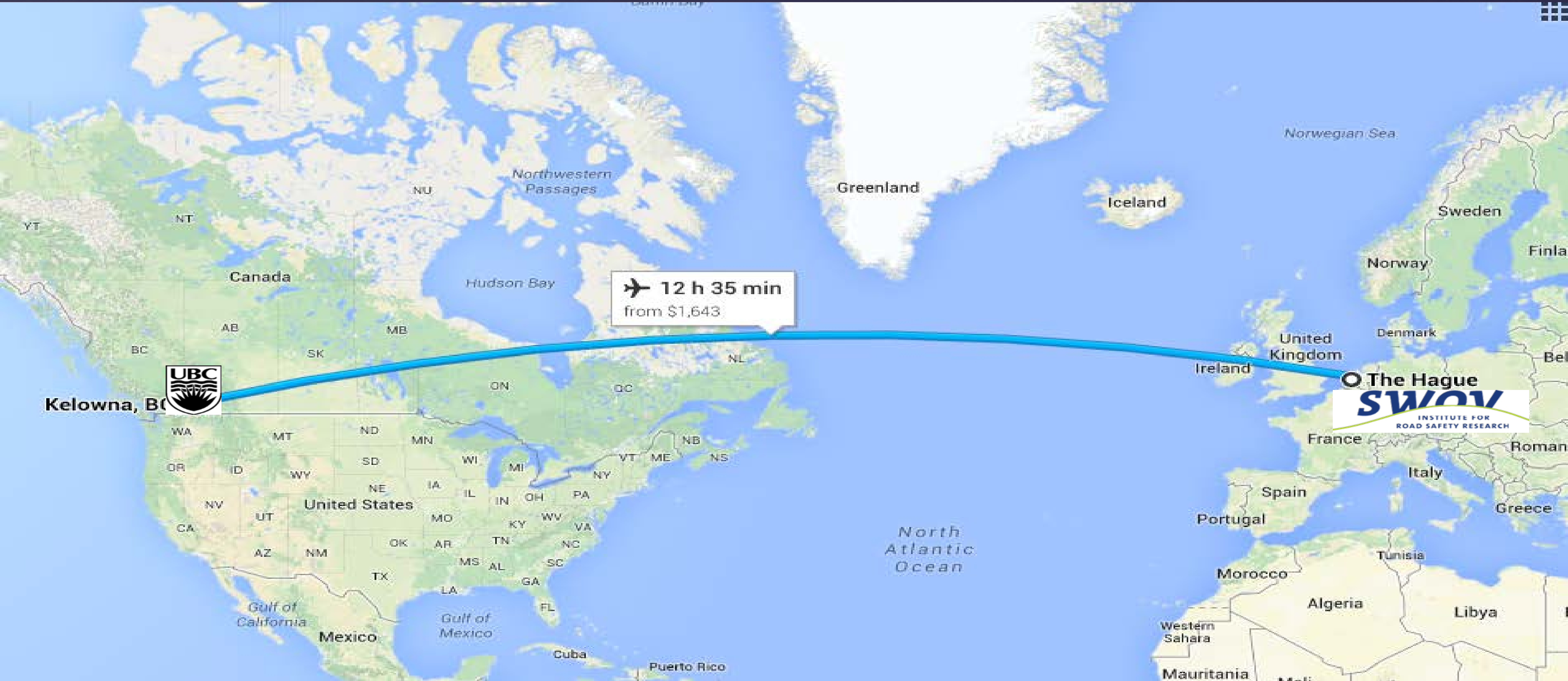
www.ubc.ca/okanagan/engineering/faculty

E-mail: gord.lovegrove@ubc.ca

Skype: lovegrove44

SWOV INSTITUTE FOR ROAD SAFETY RESEARCH

UBC Okanagan, Kelowna, BC, Canada

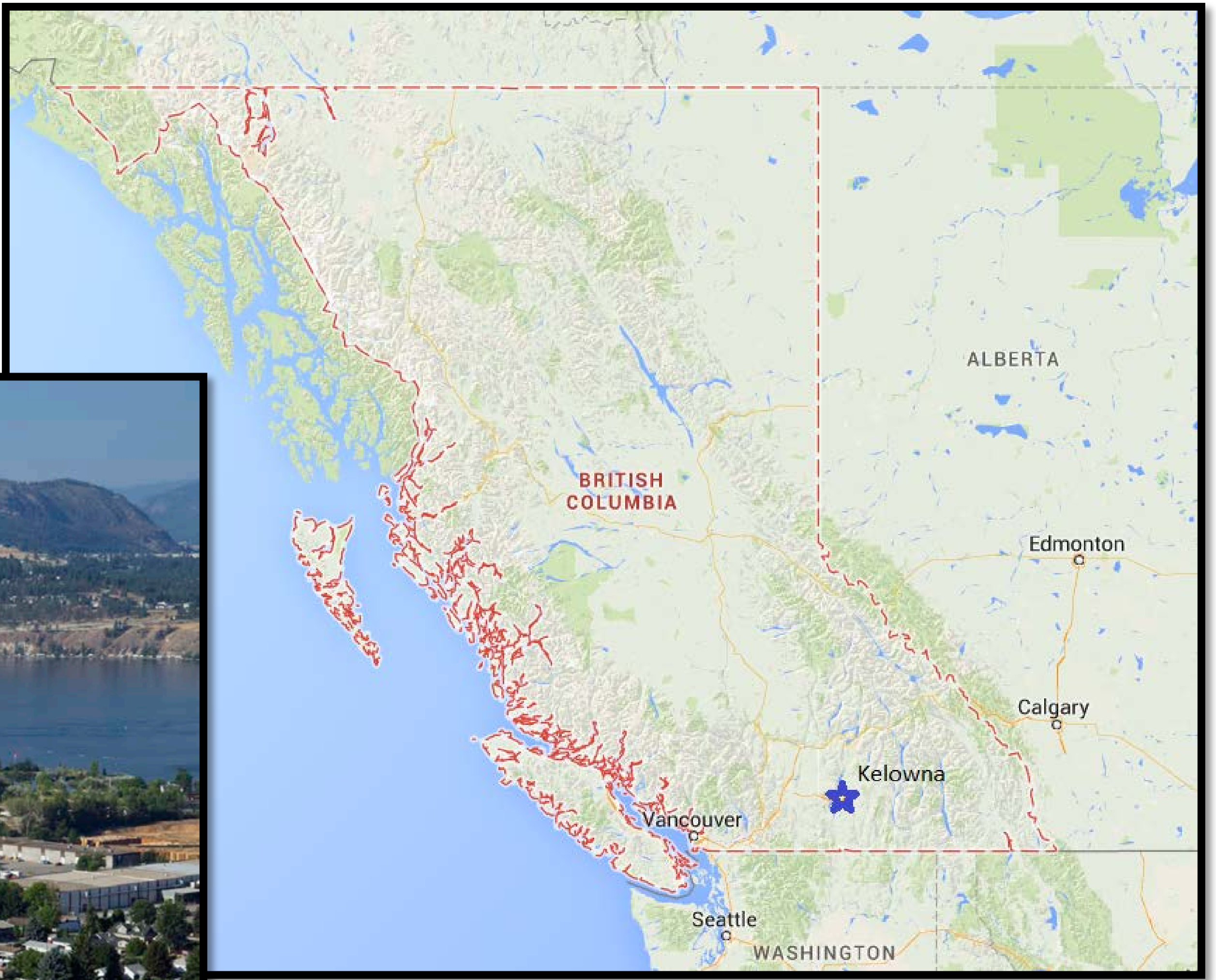


Motivation: Sustainability

- Environment: Reduce GHGs and carbon footprint, not to mention asphalt and smog.
- Social Mobility & Health: Cars are affordable to only 10% of the world's population while 80% can afford bicycles; obesity is the new nicotine
- Economic: The financial losses associated with road collisions can be as high as 6.6% of a country's GDP (5% in Canada; can someone provide this same statistic for Sweden? UK? NL?)

Kelowna, Canada
(Bike lane capital of NA on a per capital basis)

- Population (2011): 117,312
- Area: 211.82 km^2
- Density (2011): 553.8/km^2



Motivation: Barrier(s) to Cycling in NA?

	Chicago	Kelowna	Toronto	New York	Beijing	Netherlands
Private Transport	63		67	33	20	28
Cycle	1		8*	1	32	30
Walk	19			39	21	24
Rail	5		2	12	2	18
Taxi	1		1	5	4	-
Bus	11		22	10	21	-

As new bicycling infrastructure is built in NA, why aren't cyclists using it????

Hypothesis: Mental Barrier(s) to Cycling

- Main barrier to cycling is mental
- Real and perceived risk: Comfort and Safety of Cycling
- Compared to driving

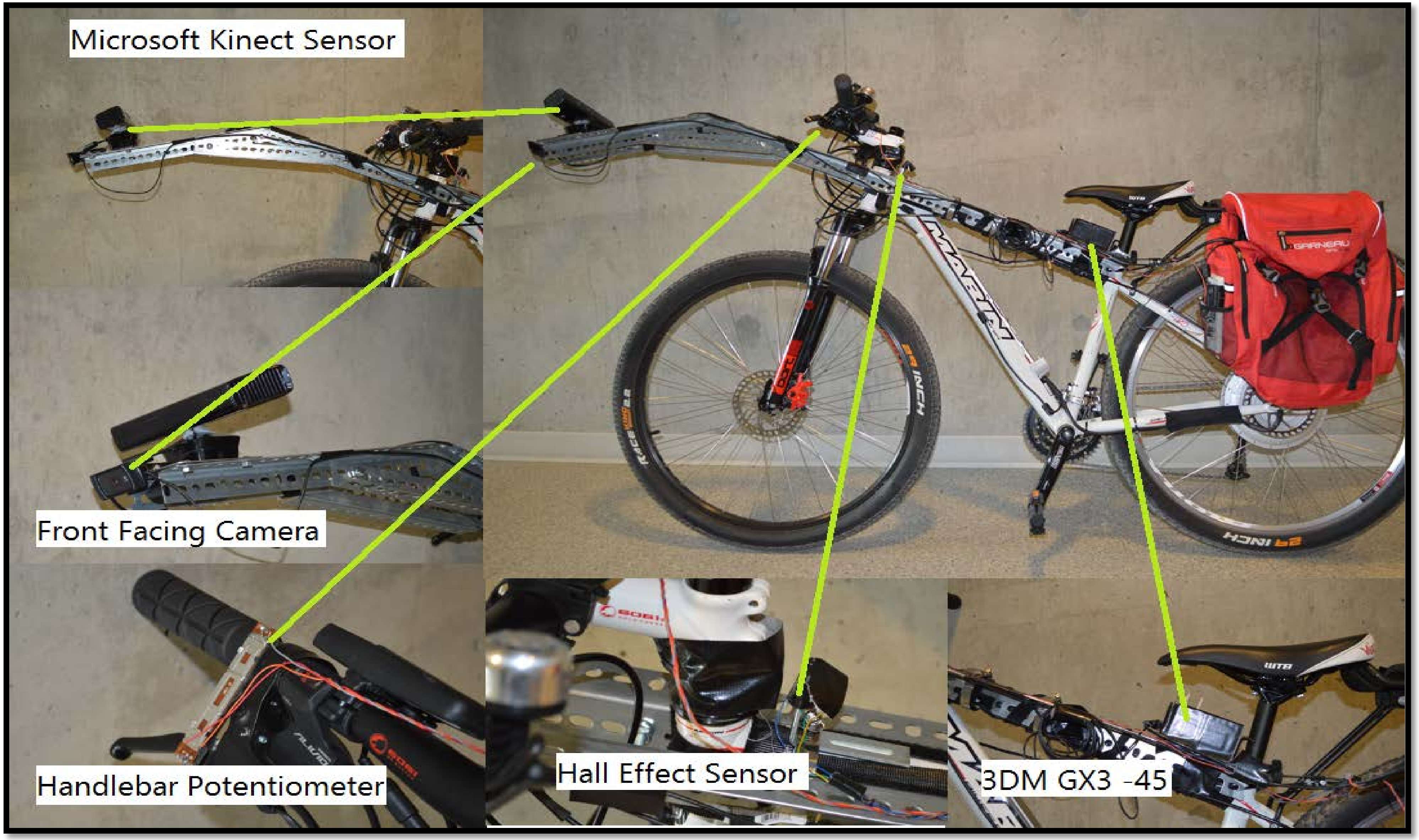
Objectives and Deliverables

1. Understand how cyclists perceive safety and comfort
2. Develop a state-of-the-art IPB to run field experiments and collect data
3. Develop a Bicycle Comfort and Safety Prediction Models (BCSPM) as a design decision-aid for engineers and planners

Methodology: Instrumented Probe Bicycle (IPB)

Sensor Type	Sensor Name	Data streams Provided
Camera	Logitech HD Pro Webcam C910	RGB Video
Time-of-Flight sensor	Microsoft Kinect 1 and 2	Depth Video, RGB Video, IR Video
Potentiometer	Hand-brake sensor (PTB6043-2010BPB103)	Hand-Brake Depression
Hall Effect Sensor	Handle-bar sensor (A1324)	Handlebar Position
GPS-Aided Inertial Navigation System	3DM GX3 -45	position, velocity (3-axis), Roll/Pitch/Yaw, Elevation

UBC STS Instrumented Probe Bicycle



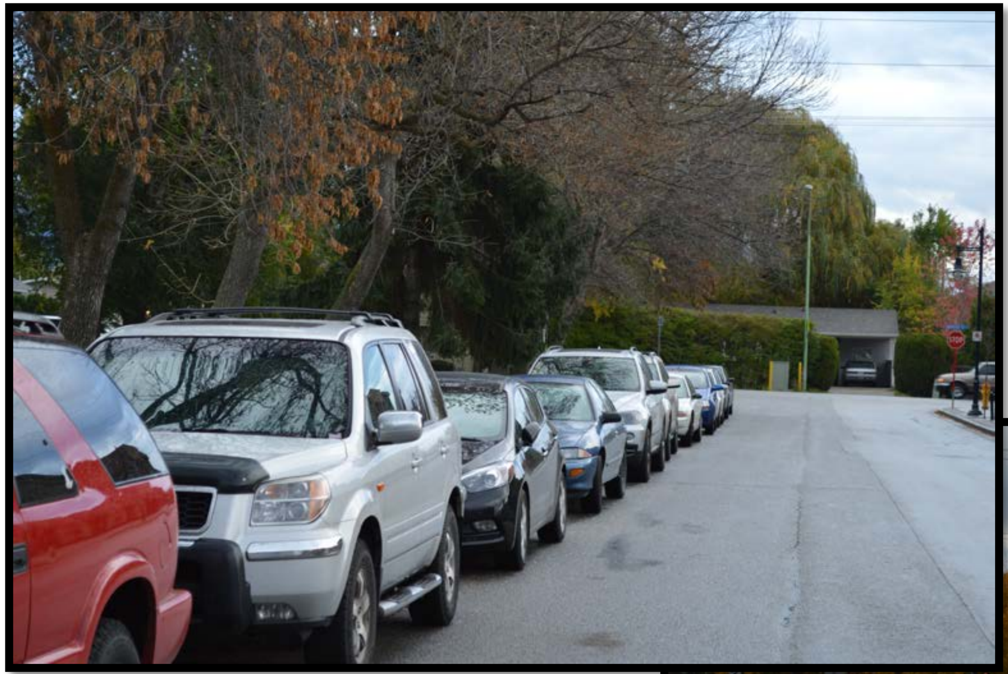
Design of Experiment

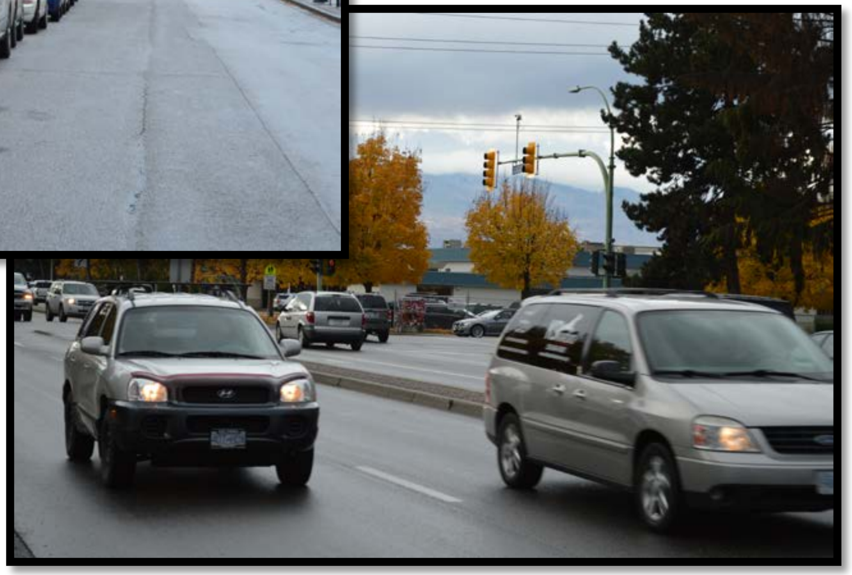
Scenario		A	B	C	D	E	F	G	H*	I*	J	K	L	M	N	O	P	Q	R	S
Road Type	Local	x			x			x			x			x			x			
	Collector		x			x			x			x			x			x		
	Arterial			x			x			x			x			x			x	
	None																			x
Bike Path Type	On Road	x	x	x							x	x	x							
	Bike Lane				x	x	x							x	x	x				
	Separated Shared Path							x	x	x							x	x	x	x
Parking	On street	x	x	x	x	x	x	x	x	x										
	None										x	x	x	x	x	x	x	x	x	x

Data

Videos were used to obtain:

- Number of parked cars
- Traffic volume
- Number of curves
- Number of obstructions
- Total Stopped Time
- Total Riding Time
- Number of Stops
- Total Travelling Time





Experiment (proof of concept)

- 7 participants from the STS Research Lab at UBC Okanagan
  - 6 male, 1 female
  - Ages 21 to 29
- Each road segment was ridden between 3 and 7 times
- 102 data points collected to test IPB technology and BCSPM methodology

Results

- Due to some technical challenges some segments did not have complete data sets from all sensors.
- Part 1 - 27 Data points with complete video data and IMU data
- Part 2 - 87 Data points with complete video data – (24 data points removed for model validation)
- Two separate analyses were carried out, one each for Part 1 and Part 2 data sets

Dependent Variables

- 5-Point Likert Scale was chosen
  - 1: extremely unsafe/uncomfortable riding experience
  - 2: unsafe/uncomfortable riding experience
  - 3: Neutral riding experience
  - 4: safe/comfortable riding experience
  - 5: extremely safe or comfortable riding experience.

Likert Scales			
SAFETY	Freq	COMFORT	Freq
Extremely Unsafe	3	Extremely Uncomfortable	1
Unsafe	10	Uncomfortable	10
Neutral	17	Neutral	11
Safe	19	Comfortable	30
Extremely Safe	38	Extremely Comfortable	35
Total	87	Total	87

Independent Variables

- 36 Independent Variables(ID) tested
- Data from sensors, field investigation, and surveys
- Notables Include:
  - Mean_LW– mean lane width (numeric)
  - Min_LW – minimum lane width (numeric)
  - Noise – noise Level (ordinal)
  - FIT – rider Fitness (ordinal)
  - P_Class – path Type (nominal)
  - Mn_C_Spd – mean cycling speed from vide (numeric)
  - Fam – segment experience(ordinal)
  - CAR_VOL – car volume (numeric)
  - CLS_VOL – close pass car volume (numeric)

Model Development

1. Data Overview
2. Categorical Principle Component Analysis (CatPCA)
   - Spearman's Correlation Coefficient
   - Independent Variables initial Selection
3. Model Fitting
   - Ordinal Logit Regression
4. Model Validation

Categorical Principle Component Analysis (CatPCA)

- Principle Component Analysis (PCA) is a data reduction technique
  - Useful to group variables into principle components when many independent variables are highly correlated
- Nonlinear PCA introduced, aka CatPCA in SPSS
  - Plays same role as PCA, but allows for variables of mixed measurement levels (i.e. Nominal, Ordinal, Interval, Ratio)

Model Formulation

$$Prob(k \leq \text{ K}) = \ 1/\{1 + \exp(-\beta_o + \sum_{i=1}^{n} \beta_i X_i)\}$$

$$Prob(k = 1) = Prob(k \leq 1)$$

$$Prob(k = 2) = Prob(k \leq 2) - Prob(k \leq 1)$$
$$Prob(k = 3) = Prob(k \leq 3) - Prob(k \leq 2)$$
$$Prob(k = 4) = Prob(k \leq 4) - Prob(k \leq 3)$$
$$Prob(k = 5) = 1 - Prob(k \leq 4)$$

- K is the safety or comfort score ranging from 1 to 5
- is the parameter estimate value corresponding with each threshold (score) for the dependent variable (safety and comfort)
- β_i is the individual parameter estimate for each independent variable
- X_i is the corresponding variable value measured from the field

CatPCA Results

- Spearman's Rank Correlation Coefficient was calculated for each independent variable
- Nine principle component groupings or dimensions
- One variable selected from each dimension for modeling safety and comfort, respectively
- All eigenvalues > 1.0

Example of a Model to predict a Rider's Perception of Safety (at 95% level of confidence)

$$P(Safety \leq 1) = \frac{1}{1 + \exp(-2.52 + 1.89 * (LN) - 0.11 * (VOL) - 1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)}$$

$$P(Safety \leq 2) = \frac{1}{1 + \exp(-4.76 + 1.89 * (LN) - 0.11 * (VOL) - 1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)}$$

$$P(Safety \leq 3) = \frac{1}{1 + \exp(-6.47 + 1.89 * (LN) - 0.11 * (VOL) - 1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)}$$

$$P(Safety \leq 4) = \frac{1}{1 + \exp(-8.13 + 1.89 * (LN) - 0.11 * (VOL) - 1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)}$$

$$P(Safety \leq 5) = 1.0$$

Sample Calculation – top row

$$P(Safety \leq 1) = \frac{1}{1 + \exp(-2.52 + 1.89 * (LN) - 0.11 * (VOL) - 1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)} = 0.010342$$

LN = 0 (Riding on Road)

Vol = 0.76 Veh/min

Noise = 1 (quiet)

Mn_C_Spd = 19.49 km/h

Fam = 1(none)

P_class = 1(on road)

Binary Code: If P_Class = 1 (on road), then put 1 if not then 0

P_class = 0 (Bike Lane)

Binary Code: If P_Class = 2 (Bike lane), then put 1 if not then 0

$$\exp(-2.52 + 1.89 * (LN) - 0.11 * (VOL) - 1.74$$

$$\frac{1}{1.74 * (Noise) + 0.03 * \left(Mn_{C_{Spd}}\right) + 0.31 * (Fam) + 8.01 * (P_{class} = 1) + 4.25 * (P_{class} = 2)}$$

Sample Calculation – all 5 rows

$P(Safety \le 1) = 0.010342$

$P(Safety \le 2) = 0.089390$

$P(Safety \le 3) = 0.351808$

$P(Safety \le 4) = 0.740563$

$P(Safety \le 5) = 1.0$

→

$Prob(k = 1) = Prob(k \le 1) = 0.010342$

$Prob(k = 2) = Prob(k \le 2) - Prob(k \le 1) = 0.079048$

$Prob(k = 3) = Prob(k \le 3) - Prob(k \le 2) = 0.262418$

$Prob(k = 4) = Prob(k \le 4) - Prob(k \le 3) = 0.388755$

$Prob(k = 5) = 1 - Prob(k \le 4) = 0.259436$

Sample Calculation – summing

Finally the predicted safety rating would be:

$$\sum_{i=1}^{5}(k_i) * Prob(k_i) = 1 * 0.010342 + 2 * 0.079048 + 3 * 0.262418 + 4 * 0.388755 + 5 * 0.259436 = 3.807895$$

Model Validation

- Validation was performed using the 24 (of 87) data points (Part 2 set) initially removed for model validation
- Within 0.23 of the actual rating for safety and within 0.95 of the actual rating for comfort
- These are encouraging results given so few (63) data points

Test Sample	Actual Safety	Model Prediction	Error	Actual Comfort	Model Prediction	Error
Average	4.21	3.98	-0.23	4.08	3.13	-0.95

Table 7. Comfort Model and Safety Model Prediction Validations

Model Validation

- Model validation reveals:
  - Participants feeling of comfort is more subjective and variable than safety
- Suggests:
  - Other independent variables (e.g. intersections or driveways) or confounding factors (e.g. speed of cars) not accounted for in the models

Findings

Significant factors affecting rider comfort and safety in this study:

- Type of bike path: on-road, bicycle lane, separated path
- Amount of space available to cyclist: Lane width, obstructions
- Cycling speed & travel speed
- Cyclists demographics: cyclists fitness, experience, training
- Traffic: volume, speed, lateral pass distance

Challenges & Limitations

1. Limited Participant Pool
2. Assumptions in variable definitions and collection
   - i.e. traffic volume not considered for cyclists on separate path
   - Parked cars deemed negligible for cyclists on separate path
3. Manoeuvrability of IPB needs improvement
4. Close follow distance by investigator may have affected comfort and safety ratings
5. Technical challenges with hardware (sensors) and software (micro-processor, integration and initiation scripts) limited data completeness
6. Not all main effects captured by the initial 36 variables
   - i.e. presence of turning vehicles waiting at intersections, large vehicles, type of intersection crossed, the infrastructure and available space to the right of the cyclist (curb, median, sidewalk, drainage swale, road shoulder etc.)

Future Work

Phase 2 of Testing

1. Expand participant pool to general public of all demographics and rider skills
2. Focus on main effects variables
3. Remove insignificant variables
4. Add additional variables
5. Adjust the way current variables are collected
   - i.e. change PAV_CON (pavement condition) from an ordinal variable to a numeric one by directly measuring vibrations
6. Finalize IPB design and configuration


GIANT
Questions?

Dr. Gord Lovegrove

Sustainable-Transport Safety-Researcher

Associate Professor, School of Engineering

University of British Columbia, Kelowna, Canada

www.ubc.ca/okanagan/engineering/faculty

E-mail: gord.lovegrove@ubc.ca

Skype: lovegrove44

SWOV INSTITUTE FOR ROAD SAFETY RESEARCH

- CatPCA was carried out on remaining independent variables
- Spearman's Rank Correlation Coefficient was calculated for each independent variable
- Yellow indicates principle component groupings
- Blue indicates the variables selected for modeling safety and comfort, respectively
- All eigenvalues > 1.0
- Example shows data from Part 2 (87 data points)

Variables	Component Loadings Dimension									Spearman's rho	
	1	2	3	4	5	6	7	8	9	SAFTY	CMFRT
Mean_LW	-.901	-.042	.177	-.082	.113	.330	-.096	.038	.021	.517**	.375**
Min_LW	-.879	-.016	.274	-.093	.096	.248	-.082	.062	-.045	.522**	.382**
P_Class	-.874	.003	.253	-.152	.023	.307	-.101	.080	-.010	.494**	.373**
Pvmt_Cdn	.766	.221	.014	-.112	-.077	.401	.236	-.134	-.129	-.412**	-.330**
Pst_Spd	.751	.405	.204	-.294	-.203	.248	-.058	-.010	-.041	-.186	-.203
R_Class	.738	.408	.378	-.214	-.133	.205	-.109	-.010	.028	-.468**	-.364**
INTD	.673	-.069	-.401	-.071	.025	.275	-.136	.160	-.101	-.238	-.191
Debris	.592	.497	.139	-.340	-.257	.210	.045	-.185	-.090	-.284*	-.226
PrkgD	.573	-.116	-.508	.215	-.213	-.056	-.298	-.068	.079	-.302*	-.188
CLS_VOL	.570	-.058	.482	.470	.343	-.032	.108	.055	.132	-.602**	-.536**
CAR_VOL	.536	-.028	.483	.435	.382	-.046	.143	-.023	.201	-.619**	-.554**
Mn_T_Spd	-.318	.627	-.009	.541	-.151	-.046	-.003	-.258	-.233	.287*	.406**
BMI	.129	.627	-.270	-.035	.552	-.114	-.136	.392	-.104	-.008	-.056
Exp	.130	.626	-.269	-.037	.552	-.116	-.135	.390	-.109	-.019	-.133
StopD	.322	-.623	.198	.317	-.287	.137	-.367	.256	-.207	-.138	-.144
StopT_P	.327	-.617	.206	.320	-.269	.136	-.370	.283	-.193	-.138	-.159
FIT	-.223	.606	-.142	.514	-.274	.103	-.163	.035	.173	.247	.324**
Fam	-.194	.476	-.228	.430	-.265	.287	-.195	.048	.263	.285*	.286*
Noise	.288	.166	.773	.190	.225	-.057	-.209	.012	.263	-.519**	-.494**
Objects	-.103	.284	.528	-.035	-.526	-.299	.174	.321	.015	-.106	.145
Mn_C_Spd	-.295	.497	.001	.649	-.113	-.088	-.040	-.133	-.385	.265*	.371**
TEMP	-.104	-.052	-.288	.513	-.327	.118	.327	.194	.325	.158	.185
F_Curve	.086	.166	-.407	-.443	-.366	-.414	-.220	.101	.293	.158	.173
DrwyD	-.226	.318	.429	-.296	-.464	-.223	.268	.337	-.137	.025	.099
F_Und	-.178	.105	-.149	-.141	.042	.780	.184	.205	.121	.103	.030
SLOPE	.318	-.285	-.343	.335	-.005	.018	.651	.222	-.135	-.225	-.216

Table 5. CatPCA and Correlation Values - Part 2 (87 Data Points)

- Mean_LW and Min_LW were summed to give LN
- CLS_VOL and CAR_VOL were summed to give VOL
- Pearson Rank Correlation, and Deviance tests for both models show values greater than 0.05.
- Positive β_i estimates (location) values indicate a positive relationship to comfort or safety
- For instance, for a 1 unit increase in VOL, the odds of “extremely safe” versus all other categories of SAFETY perception is exp(-0.11) = 0.90 times smaller
- Example shows data for Part 2 (87 data points)

Safety Model					Comfort Model				
Variables	Est.	O/R	Wald	Sig.	Variables	Est.	O/R	Wald	Sig.
Threshold									
[SAFTY = 1.00]	2.52		0.16	0.69	[CMFRT = 1.00]	-5.49		1.66	0.20
[SAFTY = 2.00]	4.76		0.58	0.45	[CMFRT = 2.00]	-1.51		0.14	0.71
[SAFTY = 3.00]	6.47		1.08	0.30	[CMFRT = 3.00]	-0.66		0.03	0.87
[SAFTY = 4.00]	8.13		1.68	0.19	[CMFRT = 4.00]	1.70		0.18	0.67
Location									
LN	1.89	6.63	3.07	0.08	LN	0.20	1.22	0.12	0.73
VOL	-0.11	0.90	2.04	0.15	VOL	-0.05	0.95	0.41	0.52
Noise	-1.74	0.18	8.00	0.00	Noise	-2.18	0.11	11.67	0.00
Mn_C_Spd	0.03	1.03	0.24	0.63	Mn_C_Spd	0.17	1.18	2.33	0.13
Fam	0.31	1.37	1.77	0.18	FIT	0.78	2.18	1.10	0.29
[P_Class=1.00]	8.01		1.85	0.17	[P_Class=1.00]	-0.78		0.05	0.83
[P_Class=2.00]	4.25		2.27	0.13	[P_Class=2.00]	-0.56		0.07	0.79
[P_Class=3.00]	0				[P_Class=3.00]	0			
Goodness-of-Fit			Pseudo R-Square		Goodness-of-Fit			Pseudo R-Square	
Test	Chi-Sqr	Sig.	Cox and Snell	.584	Test	Chi-Sqr	Sig.	Cox and Snell	.511
Pearson	174.8	1.00	Nagelkerke	0.62	Pearson	160.37	1.00	Nagelkerke	0.55
Deviance	125.4	1.00	McFadden	0.31	Deviance	114.24	1.00	McFadden	0.28