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

A. Lee[1], L. Dias[1], S. Mohanty[3], T. Carvalho[4], J. Commandeur[2], G. Lovegrove[1*]

[1] Sustainable Transport Safety Research Laboratory, School of Engineering
University of British Columbia
Kelowna, BC V1V 1V7, Canada
gord.lovegrove@ubc.ca
*corresponding author

[2]SWOV Institute for Road Safety Research, The Hague
Department of Econometrics,
VU University Amsterdam
jacques.commandeur@swov.nl

[4] Department of Civil Engineering
Federal Technological University of Paraná
Curitiba, Paraná 80230-901, Brazil
thiagocarvalhoreis@gmail.com

[3] Department of Civil Engineering
Indian Institute of Technology
Hauz Khas, New Delhi 110016, India
sudatta.mohanty.0993@gmail.com

ABSTRACT

Common deterrents to cycling in North America are the real and/or perceived concerns on the safety, comfort, and practicality of choosing cycling over other modes of transportation, concerns that may be addressed by improved cycling facilities. The challenge lies in effectively quantifying the desirability of cycling facilities to assess return on investment for bicycle infrastructure decisions. In this paper an ordinal logit regression model is proposed as a potential Bicycle Comfort and Safety Prediction Model (BCSPM) to quantitatively predict a cyclist's perceived safety and comfort. These BCSPMs were developed by conducting experiments utilizing an Instrumented Probe Bicycle (IPB). The IPB used in this study was developed using research from around the world. Many sensors were used, including: a 3DM-GX3 inertial sensor collecting time-stamped, position, velocity, and roll/yaw/pitch angles; and, a Microsoft Kinect sensor (still being operationalized) to record time-stamped eye/head positions, facial expressions, pulse, and ambient noise levels. Data for the BCSPM was collected from IPB sensors (numeric), field assessments (subjective numeric and categorical), and IPB rider questionnaires (categorical, Likert scales of comfort/safety). This paper outlines the potential applications of the BCSPM, the early modeling results of the study, the challenges faced, potential improvements that can be made to the IPB, and the next steps in this research.

Keywords: Bicycle Comfort and Safety Prediction Model (BCSPM), Instrumented Probe Bicycle (IPB), Naturalistic Bicycling Studies, Bicycle Safety, Vulnerable Road Users.

1 INTRODUCTION

Cycling has many advantages, including environmental, social and economic [1, 2, 3, 4]. In 2010, the transport sector was responsible for 22% of global CO_2 emissions [5], with nearly 60% of greenhouse gas (GHG) emissions originating from cars and light trucks [6]. A study of Latin American cities suggests that increasing bicycle trips from a 1% to 10% mode share could reduce greenhouse gas emissions by 8.4% [2]. The health benefits for cycling are also significant, increased physical activity from cycling can increase total life span by 3 to 14 months, outweighing any negative effects such as exposure to air pollution and risk to traffic accidents [3]. Economically, cars are affordable to only 10% of the world's population while 80% can afford bicycles. Bicycles can be more cost effective than driving for trip lengths less than 20 kilometers [4]. Moreover, the benefits of investing in cycling infrastructure networks are estimated to outweigh the costs by more than four-to-one [7].

In view of these social, economic, and environmental benefits, communities and transportation professionals are seeking ways to promote greater use of bicycles. However, there is a wide gap between the number of car and bicycle users in many cities. In Chicago, 63% of commuters use private transport, only 1% use bicycles [8]. Similarly, in Melbourne, 77% use private transport and only 2% use bicycles [9]. As cycling can be integrated with public transit modes such as buses and trains, the major disincentive to cycling comes from personal vehicles.

Mental barriers to cycling originate from the real and perceived concerns for its safety, comfort, and practicality relative to driving. Comparisons of exposure-based, traffic crash injury rates show that motor vehicle occupants have lower fatality rates compared to bicyclists per billion kilometers travelled [10, 11]. A lesser but still significant mental barrier is the perceived longer travel time sometimes associated with cycling relative to driving. However, an Adaptive Stated Preference (ASP) survey conducted in Minnesota concluded that commuters are willing to ride an average of 23 more minutes in order to switch from riding on a road with on-street parking to an off-road bicycle trail [12], suggesting that the travel time is not the main concern for cyclists.

To increase ridership in North America, a better understanding of perceived rider safety and comfort is needed so that more desirable bicycle facilities and effective policies can be introduced. However a there is a lack of reliable empirical tools that can evaluate planned projects and predict the level of perceived rider safety and comfort. University of British Columbia (UBC) researchers at the Sustainable Transport Safety Research Laboratory (STS) were requested by local industry with research funding to develop bicycle comfort and safety prediction models (BCSPM) to address these needs. To collect data for model development, an Instrumented Probe Bicycle (IPB) was developed in order to allow for the collection of real-time and continuous data.

The study is in its early stages, with only a small sample of field tests conducted to date. However, these tests have allowed for development of the IPB, and for proof-of-concept data for model development methodology. Therefore, the threefold purpose of this paper is to present early results:
1) A review of previous work on IPB's and bicycle safety and comfort modelling, including the potential applications of BCSPMs;
2) Progress to date on the development of the IPB, and potential improvements that can be made to the IPB; and,
3) Early UBC STS progress on BCSPM results in a North American context, including models, challenges, and next steps in this research.

2 LITERATURE REVIEW

Harkey et al. [13] identified Davis in 1987 as the first person to model bicycle safety and comfort. Davis proposed the Bicycle Safety Index Rating (BSIR), which based bicycle safety on the physical and operational features of roadways. Two safety indexes were developed, one for road segments and one for intersections. In 1994, Sorton and Walsh incorporated the perceptions of non-researchers by presenting video recordings of known road conditions to cyclists who were then asked to provide a stress level rating between 1 and 5. These stress levels were found to correlate with geometric and traffic operating conditions. In 1997, Turner et al. [14] defined the Bicycle Suitability Score (BSS) of roadways in Texas using four factors: roadway cross-section (shoulder or travel lane width), pavement surface quality, traffic volumes, and vehicle speeds. Each factor was divided into three or five ranges of values, with each range assigned a score. In 1998, Harkey et al. [13] utilized survey data and regression modelling to develop the Bicycle Compatibility Index (BCI), a comprehensive approach to quantifying the Bicycle Level of Service (BLOS). The BCI utilizes eighteen geometric and traffic related variables. Among the variables utilized are speed limit, presence of on-street parking, bicycle lane width, and driveway density. However, since data collection techniques were outdated and since there was no consideration given to rider-related variables, the models went out of practice.

More recently, Leden et al. (2000) [15] developed a risk index model by incorporating expert judgments and a Bayesian method to evaluate the safety effect of new bicycle crossings. Allen-Munley et al. (2004) [17] constructed an ordinal logistic route safety-rating model based on injury severity data from Jersey City. Petritsch et al. (2006) [17] used stepwise regression analysis to develop a safety model that predicts the relative bicycle and motor vehicle crash rates between on-street facilities and shared-use paths located next to roadways. Du et al. (2009) [18] proposed a model simulation using MATLAB/SIMULINK of an electric bicycle in order to estimate the rider's comfort. Yao et al. (2011) [19] developed a partial proportional odds model to identify the most significant contributing factors for cyclists comfort. The model was developed using comfort ratings provided by interviewing a total of 730 cyclists riding on 29 designated bicycle roadway segments in Nanjing, China. The researchers found that the geometric planning of bicycle roadways has a significant impact on rider comfort. The study found that the width of the bike path, separation from pedestrian lanes, the number of uncontrolled access points, the presence of bus stops and the land type adjacent to the bicycle path were the most significant variables. Unfortunately, these models were not developed from the collection of reliable real-time individual rider data, leading to the lack of consistency in empirical predictions of rider perceptions of comfort and safety. This, in turn stranded many bicycle infrastructure decisions in indefensible positions when compared against other, better researched budget needs.

Most recently, researchers have utilized Instrumented Probe Bicycles (IPBs) to collect better quality data on user perceptions, in efforts to build more reliable models of cyclist comfort, safety and level of service. Vanwalleghem et al. (2013) [20] was concerned with comfort estimation due to rough surfaces and proposed a vibrational comfort evaluation method. An IPB, configured with acceleration, velocity and force sensors was used to collect data for the model. The method evaluated the vibration at all the contact points of man-machine interaction, in this case, handle-bar, seat and pedals. No models have been developed to date. Twisk et al (2013) [21] studied the safety of electrical assist bicycles on the elderly, utilizing IPBs equipped with a speedometer, a GPS, a camera, an inertial measurement unit, and a potentiometer to record steer angle and steer acceleration, but no models have been produced to date. Dozza and Fernandez (2013) [22] developed one of the most well equipped IPBs to date to study bicycle dynamics and cyclist behavior. Their vision for the future is the development of models and intelligent applications to improve the safety and mobility of bicycles and cycling (i.e. curve speed warning) in the same way that similar applications have been developed for vehicles and driving.

Joo et al. (2013) [23] used an IPB to develop the Bicycle Monitoring Index (BMI), which utilizes binary logistic regression modelling and fault tree analysis. The BMI evaluates two aspects of the bicycle environment: safety and mobility, where a failure in either aspect would constitute a failure of the bicycle environment. The data was collected along a 1.4km route with 4 chosen links (2 bicycle/car links and 2 bicycle/pedestrian links). Twenty university students were chosen as volunteers and asked to rate each link as 'satisfactory' or 'unsatisfactory'. To evaluate safety, the Cycling Stability Index (CSI) was developed as shown in equation (1):

$$Cycling\ Stability\ Index\ (CSI) = Pr(STAn = 1|Xn) = \frac{exp(f(Xn,\beta)}{1+exp[f(Xn,\beta)]} \quad (1)$$

The variable STA_n represents cycling stability. $STA_n = 1$, indicates satisfactory stability while 0 represents unsatisfactory stability. The CSI provides the probability of $STA_n = 1$, which is dependent on the independent variables X_i and coefficients β. The CSI is an input in equation (2) below, where the threshold CSI value (k1) was subjectively chosen as 0.39. For mobility modelling, bicycling speed was assessed, with a threshold speed of 5 km/hr (k2) chosen since it can be considered a 'comfortable' walking speed. Thus if the bicycle environment cannot provide cyclists with a means of travelling above a 'comfortable' walking speed (5 km/hr), it has failed. From Fault Tree Analysis, final the BMI was given by equation (2):

$$BMI = \phi = 1 - \{1 - \Pr(CSI < k1)\} * \{1 - \Pr(Speed < k2)\} \quad (2)$$

The most recent and comprehensive published IPB comfort and safety research was by Yamanaka et al. (2013) [24]. The study developed one model addressing each of the five topics of interest: 'safe sense to other traffic', 'discomfort in roughness of road surface',' discomfort of narrow bicycle space', 'comfort of cycling speed' and 'total level of comfort.' To develop the models Yamanaka et al. conducted a total of 1432 IPB trials using a total of 74 street segments, 4 to 6 riders, 7 cities, and 3 countries. Riders rode each segment between 4 and 8 times, resulting in each segment being ridden between 16 and 32 times. Due to measuring system errors, from the 1432 trials, a total of 1164 samples were obtained. To evaluate the riders' sense of safety and comfort, subjects were asked to provide ratings on a five point Likert scale through a microphone as they passed each road segment for each of the topics of interest. More than 30 independent variables assessing parameters such as speed, braking, acceleration, and traffic volume were evaluated for model development. To assess pedestrian, cyclist, and vehicle density, the investigators analyzed records from a video camera mounted on the front of the bike. The number of road-users (pedestrians, cyclists and vehicles) within 10 m in front of the bicycle was counted every 4 seconds to estimate density. Depending on the bicycle lane type, vehicles beside bicycle lanes were also counted. Distances were estimated using the size of people, bicycles and cars and the final flow rate was estimated using assumed bicycle, pedestrian, and vehicle travelling speeds. The aim of Yamanaka's IPB research was to determine key factors that influence a cyclists' perception.

Factor and correlation analyses were utilized to observe the relationships between variables and between variables and the 5 topic questions. From the results of the factor analysis, variables were grouped into one of six factors named: speed stability, stop, vibration, steering, density, and braking. Variables that are grouped together from factor analysis indicated that they are highly correlated with each other. The correlation between each variable and the ratings provided for each of the five topic questions was also calculated. The variables utilized in the regression model for each topic were selected accordingly from the results of the factor analysis and correlation analysis. Since all comfort and safety scores were given on a scale of 1 to 5, the ordinal logit regression model (equation 3) was deemed appropriate and utilized in order to calculate the β values for the chosen variables.

$$Prob(score \leq \mathrm{K}) = 1/\{1 + \exp(-\beta_o + \sum_{i=1}^{n} \beta_i X_i)\} \tag{3}$$

Using the β values obtained from the regression modelling, the final models were constructed to provide a Level of Service (LOS) index for each topic of interest. Yamanaka et al. concluded that the LOS can be predicted by the speed of cycling with regard to the desired cycling speed, the standard deviation of steering angle (wobble), bicycle vertical vibration level (pavement roughness), braking behavior, traffic density in front of the cyclist, and the distance to side objects (buffer zone).

3 METHODOLOGY

3.1 Instrumented Probe Bike

Figure 1 shows the IPB used to collect data for UBC STS model development. It was a 2011 Marin Alpine Trail 29er 17" mountain bike, equipped with a BionX electric assist motor (its battery was removed and not used in this research until the IPB and BCSPM methodology are confirmed). Instruments mounted on the bicycle included a frame-mounted 3DM GX3 -45 GPS-Aided Inertial Navigation System, a straight line potentiometer mounted to the rear-wheel hand-brake lever, and a Hall effect sensor mounted onto the steering column near the handlebar, in conjunction with a rare-earth magnet. Additionally, a forward strut extension was attached to the bicycle frame in order to hold a front-facing web camera and a rider-facing Microsoft Kinect camera. The Kinect camera was mounted at a position approximately 40 cm above the front tire and 60 cm in front the handle bar. Instruments were connected via USB cable to a laptop computer which recorded data and was carried in a panier mounted on the rear rack of the bicycle. The total operating weight of the IPB was measured to be approximately 35 kg (compared to approximately 15 kg for a typical bicycle). The data streams provided by each sensor and their refresh rates are summarized in Table 1.

For each IPB field test run, the software was manually initiated before the rider could proceed onto the selected pathway. The data provided was time stamped for all sensors, other than the front-facing camera. In order to synchronize the video with the rest of the data, a visual signal was displayed in front of the camera in conjunction with the initialization of one of the other sensors, to serve as a reference point in processing.

Table 1. Summary of sensors mounted on IPB

Sensor Type	Sensor Name	Data streams Provided
Camera	Logitech HD Pro Webcam C910	RGB Video
Time-of-Flight sensor	Microsoft Kinect	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	GPS position, NED velocity, Roll/Pitch/Yaw, Elevation

3.2 Road Segments

Table 2 outlines 19 scenarios of interest, labelled A through S, based on varying three types of road conditions: road hierarchy, bike path type, and on-street parking. Street segments were selected from within Kelowna, BC to cover the 18 theoretical scenarios. Based on other research methodologies, the length of each road segment ranged from a minimum of 300 to 800 meters [24]. Each IPB test run performed on a given street segment represents a single data point or sample for modelling purposes. In some situations, multiple road segments were selected for a single scenario. Individual road segments were grouped into test routes consisting of one or two road segments. In this way, multiple segments could be tested immediately preceding (or following) each other on the ride [21, 24]. A total of 26 segments were utilized.

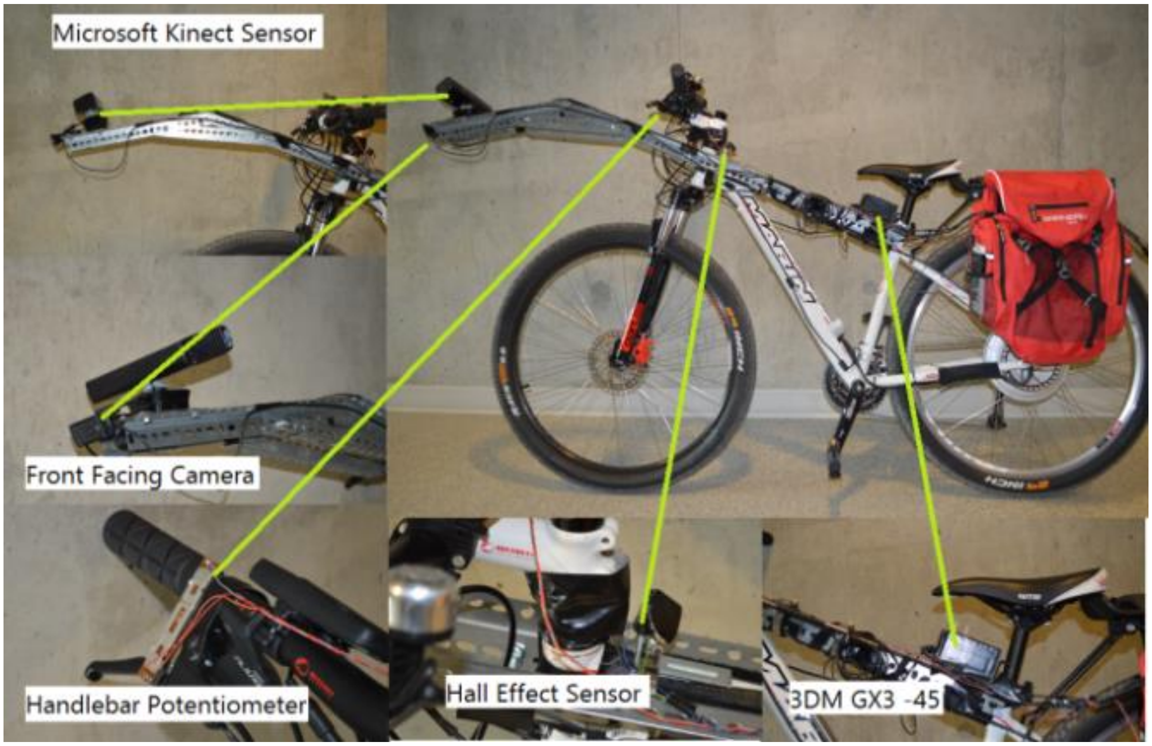

Figure 1. IPB and Mounted Sensors

Table 2.Summary of test scenarios

Scenario	Road Hierarchy			Bike Path Type			On Street Parking	
	Local	Collector	Arterial	On Road	Bike Lane	Separated Shared	Yes	No
A	x			X			x	
B		x		X			x	
C			x	X			x	
D	x				x		x	
E		x			x		x	
F			x		x		x	
G	x					x	x	
H*		x				x	x	
I*			x			x	x	
J	x			X				x
K		x		X				x
L			x	X				x
M	x				x			x
N		x			x			x
O			x		x			x
P	x					x		x
Q		x				x		x
R			x			x		x
S	NA					x		x

*No segments were found for scenarios H and I

3.3 Participant Sample Size

For these proof-of-concept test rides, participants were researchers from the UBC STS research lab. Seven test riders were involved in this study; all between the ages of 21 and 29, consisting of 6 males and 1 female. Each of the route segments was ridden between 3 and 7 times no more than once by each rider. In total 26 road segments were examined between 3 and 7 times each to account for a total 102 samples. Admittedly, a much larger sample size was required; however, in the time available over 100 samples was reasonable to test IPB technology and BCSPM methodology.

3.4 Test Data

Data for over 30 independent variables were successfully collected by the IPB, rider surveys, and field assessments. Table 3 provides definitions for all 36 variables for which initial data were collected and tested in model development.

Table 3. Variable Definitions for comfort and safety data collected for BCSPM development

#	Variable Name	Units	Description
1	T_SPD_V	km/h	Mean Travel Speed. Mean travel speed describes the speed at which the segment was traversed using travel time which indicates the total time elapsed for traveling a road segment. Calculated from front-mounted video camera
2	C_SPD_V	km/h	Mean cycling speed describes the speed at which the segment was traversed using cycling time, where cycling time describes the time spent in motion. Calculated from front-mounted video camera
3	R_STP	sec/sec	Total stopped time divided by travel time. Stopped time describes the amount of time spent stationary (waiting for red-light, etc.).
4	F_STP	#/km	Number of stops divided by segment length.
5	NOIS	Indicator	Environmental noise (classified based on the judgment of the investigator as “quiet”, “medium” or “loud”).
6	CAR_SPD	Indicator	Speed limit of the road (classified as “0 km/h”, “40 km/h”, “50 km/h”, “60 km/h”, or “70 km/h”) A speed limit of 0 km/h are assigned to bike paths not adjacent to a road.
7	CAR_VOL	Veh/min	Perceived vehicle volume travelling in the same direction as the participant (applies only to participants cycling on road or within a bike lane). This parameter is calculated by counting the number of vehicles to pass the participant (in the same direction), divide by cycling time.
8	CLS_VOL	Veh/min	Perceived close vehicle pass volume travelling in the same direction as the participant. Vehicles passing the participant within the same lane or within the immediately adjacent lane are considered a close pass. This parameter is calculated by counting the number of vehicles to pass the participant and dividing that value by the cycling time. (Applies only to participants cycling on road or within a bike lane).
9	PTYPE	Indicator	Path type (classified as “on road”, “bike lane”, and “separated path”).
10	F_OBS	#/km	Obstruction frequency (per kilometer of the length of the segment). Obstructions are defined as objects placed within the participants cycling path that causes the participant to execute a maneuver in order to avoid it.
11	SLOP	%	Average slope.
12	PAV_CON	Indicator	Pavement condition (classified as “very poor”, “poor”, “fair”, “good” or “very good”). Classifications are based on the Pavement Condition Rating outlined within the FHWA Highway Performance Monitoring

			System [14]
13	R_MAIN	Indicator	Road maintenance describes the amount of debris such as sticks, leaves, sand or gravel on the path (classified based on the judgment of the investigator as “poor”, “fair”, “good”, or “very good”) A “very good” rating indicates no debris on the path, rating of “good” indicates that there is some debris on path but it is minor and easily overlooked. A “Fair” rating indicates that there are noticeable patches of debris that may cause some discomfort. Finally, a “poor” rating indicates that there is an abundant amount of debris on the path that significantly disturbs the riding experience.
14	F_UND	#/km	Undulations frequency (per kilometer of the length of the segment). Undulations are defined as a sudden rise or drop of 20 cm in height or more.
15	MIN_LN	m	Minimum lane width of the bike path in the road segment (a lane width of 0 is assigned when riding on the road). On-street bicycle lanes were assumed to have the minimum specified width of 1.5 m.
16	MAJ_LN	m	Lane width of the bike path in the majority of the road segment (a lane width of 0 is assigned when riding on the road). On-street bicycle lanes were assumed to have the minimum specified width of 1.5 m.
17	F_CURV	#/km	Curve frequency (per kilometer of the length of the segment). Based on the judgment of the investigator, a curve was loosely defined within this study as any non-gradual change in the alignment of the bike path.
18	F_INS	#/km	Intersection frequency (per kilometer of the length of the segment). Intersections frequency describes the number of intersections that the participant crossed during the riding of each segment.
19	TEMP	Indicator	Temperature (classified as “below 0 °C”, “0-10 °C”, “10-20 °C”, “20-30 °C” or “Above 30°C”).
20	HIER	Indicator	Road hierarchy (“none” in the case where no roads are adjacent to the bike path, “local”, “collector” or “arterial”).
21	WIND	Indicator	Wind strength (classified based on the judgment of the investigator as “none/light”, “medium” and “strong”).
22	PRK_CAR	#/km	Parked car density (the number of parked cars per kilometer of length of the segment).
23	ADJ_DRWY	#/km	Adjacent driveway density (per kilometer of the length of the segment, includes parking lot accesses and alley ways).
24	AGE	Indicator	Age (“19-29 yrs”, “30-39 yrs”, “50-59 yrs” or “over 60 yrs”)
25	FIT	Indicator	Fitness is described as hours of exercise per week (“less than 2 hrs”, “2 to 5 hrs”, “5 to 10 hrs” or “greater than 10 hrs”).
26	GEN	Indicator	Gender (classified as “male” or “female”).
27	R_EXP	Indicator	Riding experience is described by the length of time of which the participant has known how to ride a bicycle (classified as “less than 1 year”, “1 to 5 years”, “5 to 10 years” or “greater than 10 years”).
28	BMI	BMI	Body mass index (BMI)
29	S_EXP	Indicator	Segment experience describes the number of times the participant has previously ridden on the segment (classified as “none”,”1 time”,”1-3 times”, “3-5 times”, or “more than 5 times”).
30	ROL_SDEV	Degrees	The standard deviation of the rotation about the roll axis (axis about the direction of travel of a bicycle). Calculated from Inertial sensor data.
31	PCH_SDEV	Degrees	The standard deviation of the rotation about the pitch axis (axis parallel to the wheel axles). It is hypothesized that pitch can be used to infer surface roughness. Calculated from Inertial sensor data.
32	T_SPD_INS	Km/h	Mean Travel Speed. Calculated from Inertial sensor data.
33	T_SPD_SDEV	Km/h	The standard deviation of mean travel speed. Calculated from Inertial sensor data.

34	M_SPD	Km/h	Maximum cycling speed
35	C_SPD_INS	Km/h	Mean Cycling Speed. Calculated from Inertial sensor data.
36	C_SPD_SDEV	Km/h	The standard deviation of mean cycling speed. Calculated from Inertial sensor data.

3.5 *Data Collection*

Upon completion of the test rides, participants provided a rating for the safety and comfort level they experienced during each road segment. These ratings were provided based on a Likert scale from 1 to 5, with 1 representing an extremely unsafe or uncomfortable riding experience to 5 representing an extremely safe or comfortable riding experience. A rating of 3 represents a neutral experience. Cycling safety was defined as the risk of physical or psychological damages due to cycling accidents. Cycling comfort was defined as the physical or psychological ease and convenience of cycling. The sense of safety can be summarized to be the feeling of potential danger, while comfort can be summarized to be the feeling of enjoyment. Participants were asked about their previous experience with the ridden segments, personal characteristics (age, fitness, BMI, gender), and riding experience. Other route data - pavement condition, road maintenance, temperature, wind strength, environmental noise, speed limit and lane widths - were recorded separately by the investigators. Video data collected from the forward facing camera on the IPB was post-processed to obtain the remaining variables associated with traffic volumes, obstructions, curves, parked cars, intersections and driveways. Google Earth maps were also utilized to assist with the counting of intersections and driveways. Finally, data collected by the GPS-Aided Inertial Navigation System (3DM-GX3-45) was post-processed to obtain real-time bicycle orientation and velocity data.

From the 102 samples collected, video recording errors were observed for 15 samples. Additional operational and technical challenges with the GPS-Aided Inertial Navigation System (3DM-GX3-45), Microsoft Kinect sensor, handle-bar sensor and hand-brake sensor resulted in discontinuous data and missing data for many more samples. As a result of these challenges samples often did not contain a complete set of data from all sensors. These challenges were noted for future work and preliminary BCSPM were developed from two separate analyses (Part 1 and Part 2) utilizing data collected from the 3DM-GX3-45 and the webcam. Of the 102 samples collected, a total of 27 samples contained complete webcam and the 3DM-GX3-45 data. These 27 samples were utilized in Part 1. A total of 87 samples contained complete webcam data; these samples were utilized in Part 2. The model developed in Part 2 used 63 of those data points (roughly 75% of the 87 data points) referred to as model data. The remaining 24 data points were chosen randomly and kept to validate the developed model.

Table 4 summarizes the continuous variable data. Table 5 summarizes indicator variable data, and Table 6 summarizes dependent variable data. Although more IPB test runs using non-staff participants would have been ideal (and will be carried out in phase 2), this was sufficient data to run model development using SPSS software and Categorical Principle Component Analysis (CatPCA). Future research (Phase 2 Spring 2015) would then refine and use this methodology with a substantially larger and complete data set.

4 MODEL DEVELOPMENT AND RESULTS

4.1 Principle Component Analysis (PCA) and Correlation Analysis

First, a brief assessment of the data resulted in the removal of 3 variables from further analysis. Gender variables were removed from the analysis as the demographics consisted of 6 male participants and 1 female participant which is inadequate for measuring gender differences. WIND and AGE were also removed as they remained constant throughout data collection.

Table 4. Numeric (Continuous) Data Summary Part 1 / 2 (Values for 27 / 87 Data Points)

Variables	Units of Measure	Minimum	Maximum	Mean	Std. Deviation
T_SPD_V	Km/h	6.53 / 6.53	27.43 / 27.43	15.81 / 15.84	4.10 / 4.29
C_SPD_V	km/h	7.83 / 7.83	41.74 / 41.74	16.60 / 16.54	5.97 / 5.01
R_STP	sec/sec	0.00 / 0.00	0.34 / 0.35	0.03 / 0.04	0.08 / 0.09
F_STP	#/km	0.00 / 0.00	8.89 / 8.89	0.66 / 0.52	1.84 / 1.35
CAR_VOL	Veh/min	0.00 / 0.00	14.88 / 14.88	1.96 / 1.92	3.25 / 3.17
CLS_VOL	Veh/min	0.00 / 0.00	5.08 / 8.37	1.32 / 1.37	1.69 / 2.07
F_OBS	#/km	0.00 / 0.00	6.67 / 6.67	1.00 / 0.74	2.05 / 1.81
SLOP	%	-1.38 / -1.38	1.13 / 1.25	-0.06 / -0.02	0.60 / 0.40
F_UND	#/km	0.00 / 0.00	9.23 / 9.23	0.68 / 0.59	2.46 / 1.85
MIN_LN	m	0.00 / 0.00	3.80 / 3.80	1.48 / 1.37	1.18 / 1.29
MAJ_LN	m	0.00 / 0.00	3.80 / 3.80	1.54 / 1.53	1.23 / 1.44
F_CURV	#/km	0.00 / 0.00	6.67 / 6.67	1.24 / 0.72	2.26 / 1.78
F_INS	#/km	0.00 / 0.00	12.00 / 12.00	4.89 / 4.70	4.14 / 3.66
PRK_CAR	#/km	0.00 / 0.00	88.89 / 125.71	12.02 / 19.93	21.51 / 32.22
ADJ_DRWY	#/km	0.00 / 0.00	46.67 / 46.67	20.58 / 13.49	17.00 / 16.89
BMI	BMI	20.28 / 20.28	28.59 / 31.75	24.82 / 24.51	3.02 / 3.06
ROL_SDEV*	Degrees	1.39	7.47	2.68	1.45
PCH_SDEV*	Degrees	0.67	1.73	1.14	0.28
T_SPD_INS*	Km/h	6.77	22.94	15.08	3.64
T_SPD_SDEV*	Km/h	1.55	6.96	3.82	1.47
M_SPD*	Km/h	13.10	29.12	20.52	4.24
C_SPD_INS*	Km/h	7.52	22.94	15.37	3.47
C_SPD_SDEV*	Km/h	1.55	8.60	3.70	1.58

*Variables only available for Part 1 (27 data points)

Categorical Principle Component Analysis (CatPCA) was then performed in order to observe the relationship between the remaining independent variables. As noted, using CatPCA allows logit regression analysis to be conducted in SPSS without the need for normality, so long as variables are fully disclosed as either Scale (e.g. Speed), Ordinal (e.g. Likert Scale), or Nominal (e.g. Gender). It should be noted that path type (PTYPE) and road hierarchy (HIER) were treated as ordinal variables rather than nominal throughout the preliminary analysis as categories within each variable can be ranked. Tables 10 and 11 contain the results of the component loading dimensions from the CatPCA and Spearman correlation analyses between each of the independent variables and the safety and comfort ratings. The results reveal 7 dimensions or latent variables. The correlation target level of confidence of 95% was desired, in some cases 85% was accepted (Part 1), within Part 2 many cases exceeded 99%.

The cells highlighted in Table 7 and Table 8 under component loading dimensions indicates variables with the highest correlation to their respective dimensions. In future model development, all well correlated (i.e. 95% or more) variables in each factor could be combined and the model re-run. For example, in part 2, dimension 1 we could introduce a new variable in SPSS called PATH=MAJ_LN + MIN_LN + PTYPE, because all variables are significantly correlated but collinear, thus avoiding the problem of collinear model variables. For each case of the analysis (Part 1 and Part 2) the most important component dimensions are discussed.

Table 5. Numeric (continuous) data summary Part 1 / 2 (Values for 27 / 87 Data Points)

NOIS	Frequency	Percent
Quiet	12 / 46	44.4 / 52.9
Medium	14 / 35	51.9 / 40.2
Loud	1 / 6	3.7 / 6.9
Total	27 / 87	100 / 100
FIT	Frequency	Percent
2-5 hrs	15 / 48	55.6 / 55.2
5-10 hrs	12 / 39	44.4 / 44.8
Total	27 / 87	100 / 100
PTYPE	Frequency	Percent
On Road	8 / 35	29.6 / 40.2
Bike Lane	10 / 20	37.0 / 23.0
Separated Path	9 / 32	33.3 / 36.8
Total	27 / 87	100 / 100
PAV_CON	Frequency	Percent
Fair	3 / 12	11.1 / 13.8
Good	11 / 26	40.7 / 29.9
Very Good	13 / 49	48.1 / 56.3
Total	27 / 87	100 / 100
R_MAIN	Frequency	Percent
Fair	0 / 10	0 / 11.5
Good	5 / 11	18.5 / 12.6
Very Good	22 / 66	81.5 / 75.9
Total	27 / 87	100 / 100
AGE	Frequency	Percent
19-29	27 / 87	100 / 100
WIND	Frequency	Percent
None/Light	27 / 87	100 / 100

HIER	Frequency	Percent
none	0 / 8	0 / 9.2
Local	16 / 42	59.3 / 48.3
Collector	2 / 11	7.4 / 12.6
Arterial	9 / 26	33.3 / 29.9
Total	27 / 87	100 / 100
CAR_SPD	Frequency	Percent
0 km/h	0 / 8	0 / 9.2
40 km/h	4 / 8	14.8 / 9.2
50 km/h	19 / 63	70.4 / 72.4
60 km/h	4 / 8	14.8 / 9.2
Total	27 / 87	100 / 100
R_EXP	Frequency	Percent
less than 1 year	2 / 6	7.4 / 6.9
10+ years	25 / 81	92.6 / 93.1
Total	27 / 87	100 / 100
TEMP	Frequency	Percent
11-20 C	6 / 17	22.2 / 19.5
21-30 C	21 / 70	77.8 / 80.5
Total	27 / 87	100 / 100
GEN	Frequency	Percent
Male	18 / 65	66.7 / 74.7
Female	9 / 22	33.3 / 25.3
Total	27 / 87	100 / 100
S_EXP	Frequency	Percent
none	19 / 60	70.4 / 69.0
once	6 / 13	22.2 / 14.9
1 to 3 times	2 / 2	7.4 / 2.3
Total	27 / 87	100 / 100

In part 1, dimension 1 includes variables that can be related in some way to mobility and the physical cycling environment. T_SPD_V, C_SPD_V, C_SPD_INS, M_SPD and T_SPD_INS are all variables that describe speed, while F_OBS and ADJ_DRWY both describe situations where the cyclist would be inclined to slow down to watch for other road users or to maneuver around an obstacle as shown by the variable ROL_SDEV. Further, F_STP directly describes the frequency of stops during the ride and R_EXP and BMI are both rider characteristics that can greatly affect their mobility and speed through a segment. The second dimension includes variables such as SLOP, PAV_CON, F_UND, PTYPE, MAJ_LN and MIN_LN. All of these variables describe the physical features of cycling path. Dimension 3 contains variables that describe the stability of the bicycle and speed (PCH_SDEV, T_SPD_SDEV, R_STP and C_SPD_SDEV) and dimension 4 contains variables that are related to traffic volume (HIER, CLS_VOL, CAR_VOL and NOIS).

Table 6. Dependent variable data summary

Part 1 (27 Data Points)			**Part 2 (87 Data Points)**		
SAFETY	**Frequency**	**Percent**	**SAFETY**	**Frequency**	**Percent**
Extremely Unsafe	0	0	Extremely Unsafe	3	3.4
Unsafe	2	7.4	Unsafe	10	11.5
Neutral	5	18.5	Neutral	17	19.5
Safe	7	25.9	Safe	19	21.8
Extremely Safe	13	48.1	Extremely Safe	38	43.7
Total	27	100.0	Total	87	100.0
CMFRT	**Frequency**	**Percent**	**CMFRT**	**Frequency**	**Percent**
Extremely Uncomfortable	0	0	Extremely Uncomfortable	1	1.1
Uncomfortable	1	3.7	Uncomfortable	10	11.5
Neutral	6	22.2	Neutral	11	12.6
Comfortable	7	25.9	Comfortable	30	34.5
Extremely Comfortable	13	48.1	Extremely Comfortable	35	40.2
Total	27	100.0	Total	87	100.0

Within Part 2 the majority of variables are split between dimension 1 and dimension 2. Dimension 1 includes variables that describe traffic volume and speed (CAR_SPD, CLS_VOL, CAR_VOL), and the type of corridor (MAJ_LN, MIN_LN, PTYPE, R_MAIN, PAV_CON, PRK_CAR , HIER and F_INS.) Dimension 2 contains variables that are associated directly (T_SPD_V, R_STP and F_STP) or indirectly (FIT, BMI, R_EXP and S_EXP) with cycling speed and speed stability.

Variables were selected for modelling based on their correlation with SAFTY and CMFRT. Models were initially constructed utilizing all the variables marked with an asterisk within the Spearman's rho correlation column in Table 7 and Table 8. Variables were then subsequently removed based on the appropriateness of the sign of the resulting β value estimates, the sig. value of the β value estimates, and in some cases to achieve an acceptable goodness-of-fit sig. value above 0.05. The highlighted

Spearman's values correspond with the variables used in the final model. As MAJ_LN and MIN_LN as well as CLS_VOL and CAR_VOL were highly collinear, they were summed into variables LN and VOL respectively for modelling purposes. It should be noted that PAV_CON and R_MAIN is significantly correlated with safety and comfort, however, in a negative correlation which is not intuitive. This is likely due to the positive correlation that PAV_CON and R_MAIN has with CAR_SPD, CLS_VOL and CAR_VOL over powering its intuitively positive effect on comfort and safety. It can be reasoned that roads with more traffic volume and higher speeds are more frequently maintained. Additionally, it was observed during test riding that shared paths were often less well maintained than bike lanes or roads.

Table 7. Component Loadings and Correlation Values - Part 1 (27 Data Points)

Variables	Component Loadings Dimension									Spearman's rho	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**SAFTY**	**CMFRT**
T_SPD_V	.896	-.085	-.170	.243	.163	-.056	.020	-.199	-.039	-.100	.283
C_SPD_V	.883	-.076	-.200	.246	.168	-.046	.037	-.189	-.080	-.074	.282
BMI	.872	-.144	-.092	-.081	-.232	.191	.011	.124	-.204	.231	.166
C_SPD_INS	.871	-.108	-.195	.345	.158	-.151	.053	-.047	.015	-.043	.312*
T_SPD_INS	.871	-.111	-.276	.257	.166	-.196	.010	-.048	.018	-.063	.218
R_EXP	.861	-.152	-.124	-.081	-.184	.174	.039	.125	-.229	.303*	.245
F_STP	-.792	.122	.481	-.024	-.113	.127	.069	.227	-.081	.106	.036
ROL_SDEV	-.682	-.335	.278	-.020	.246	-.118	-.126	-.261	.248	.045	-.071
ADJ_DRWY	-.671	.086	-.105	.290	.315	-.487	.161	.032	-.058	0.000	.104
F_OBS	-.669	.086	-.099	.293	.321	-.485	.161	.036	-.058	-.405**	-.026
M_SPD	.615	-.537	.095	.241	-.042	-.224	-.004	.299	.034	-.066	.359*
SLOP	.004	.859	-.039	.373	-.233	-.093	.036	.065	-.156	-.135	-.205
PAV_CON	.014	.806	.009	.401	-.316	-.142	.007	.175	-.099	-.205	-.277
F_UND	-.046	-.566	-.181	-.504	.248	.224	.145	.089	.289	.078	-.049
PTYPE	-.435	-.550	-.465	.110	-.381	-.098	.226	-.235	-.107	.210	.283
MAJ_LN	-.434	-.550	-.466	.112	-.381	-.098	.225	-.235	-.109	.266	.335*
MIN_LN	-.433	-.547	-.467	.112	-.383	-.098	.225	-.237	-.108	.271	.319*
C_SPD_SDEV	-.072	-.471	.694	.354	-.070	-.025	-.253	-.037	-.044	-.310*	.083
T_SPD_SDEV	.002	-.350	.681	.388	-.073	.278	-.019	-.323	-.169	-.239	.150
PCH_SDEV	.189	-.515	.666	.138	-.170	-.198	-.112	-.084	.073	-.161	.131
PRK_CAR	.288	.418	.600	.066	.321	.144	.206	-.232	-.046	-.033	-.040
R_STP	-.435	.132	.568	.205	-.137	.465	.229	-.091	-.175	.151	.054
S_EXP	.352	-.241	.549	.431	-.162	.036	.316	-.104	.234	.054	.056
HIER	-.336	-.303	.068	.696	.139	-.249	-.325	.217	-.151	-.394**	.025
F_CURV	.193	-.337	.200	-.608	.555	.095	-.058	-.007	-.083	.151	.180
CLS_VOL	-.068	.298	-.388	.575	.024	.421	-.214	-.238	.330	-.297*	-.436**
FIT	.403	-.118	.365	.573	.179	-.212	.346	.137	.107	-.103	.217
CAR_VOL	-.069	.290	-.406	.567	-.016	.416	-.280	-.215	.286	-.315*	-.401**

NOIS	-.318	-.473	-.246	.486	.221	.057	-.335	.334	.134	-.424**	-.287*
R_MAIN	.085	-.339	.113	.061	-.563	.283	.041	.486	.221	.138	-.112
CAR_SPD	-.263	-.211	-.208	.244	.534	.477	.010	.075	-.458	.146	.188
F_INS	.337	.364	.331	-.422	-.217	-.473	-.087	-.175	.204	-.053	-.183
TEMP	-.114	.120	-.139	.231	.312	.236	.723	.216	.269	.166	-.025

**Correlation is significant at the 0.05 level (i.e. 95% level of confidence)

* Correlation is significant at the 0.15 level (i.e. 85% level of confidence)

Table 8. Component Loadings and Correlation Values - Part 2 (87 Data Points – 75% of data)

Variables	Component Loadings Dimension									Spearman's rho	
	1	2	3	4	5	6	7	8	9	SAFTY	CMFRT
MAJ_LN	-.901	-.042	.177	-.082	.113	.330	-.096	.038	.021	.517**	.375**
MIN_LN	-.879	-.016	.274	-.093	.096	.248	-.082	.062	-.045	.522**	.382**
PTYPE	-.874	.003	.253	-.152	.023	.307	-.101	.080	-.010	.494**	.373**
PAV_CON	.766	.221	.014	-.112	-.077	.401	.236	-.134	-.129	-.412**	-.330**
CAR_SPD	.751	.405	.204	-.294	-.203	.248	-.058	-.010	-.041	-.186	-.203
HIER	.738	.408	.378	-.214	-.133	.205	-.109	-.010	.028	-.468**	-.364**
F_INS	.673	-.069	-.401	-.071	.025	.275	-.136	.160	-.101	-.238	-.191
R_MAIN	.592	.497	.139	-.340	-.257	.210	.045	-.185	-.090	-.284*	-.226
PRK_CAR	.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**
T_SPD_V	-.318	.627	-.009	.541	-.151	-.046	-.003	-.258	-.233	.287*	.406**
BMI	.129	.627	-.270	-.035	.552	-.114	-.136	.392	-.104	-.008	-.056
R_EXP	.130	.626	-.269	-.037	.552	-.116	-.135	.390	-.109	-.019	-.133
F_STP	.322	-.623	.198	.317	-.287	.137	-.367	.256	-.207	-.138	-.144
R_STP	.327	-.617	.206	.320	-.269	.136	-.370	.283	-.193	-.138	-.159
FIT	-.223	.606	-.142	.514	-.274	.103	-.163	.035	.173	.247	.324**
S_EXP	-.194	.476	-.228	.430	-.265	.287	-.195	.048	.263	.285*	.286*
NOIS	.288	.166	.773	.190	.225	-.057	-.209	.012	.263	-.519**	-.494**
F_OBS	-.103	.284	.528	-.035	-.526	-.299	.174	.321	.015	-.106	.145
C_SPD_V	-.295	.497	.001	.649	-.113	-.088	-.040	-.133	-.385	.265*	.371**
TEMP	-.104	-.052	-.288	.513	-.327	.118	.327	.194	.325	.158	.185
F_CURV	.086	.166	-.407	-.443	-.366	-.414	-.220	.101	.293	.158	.173
ADJ_DRWY	-.226	.318	.429	-.296	-.464	-.223	.268	.337	-.137	.025	.099
F_UND	-.178	.105	-.149	-.141	.042	.780	.184	.205	.121	.103	.030
SLOP	.318	-.285	-.343	.335	-.005	.018	.651	.222	-.135	-.225	-.216

****Correlation is significant at the 0.01 level (i.e.99% level of confidence)**

*Correlation is significant at the 0.05 level (i.e. 95% level of confidence)

4.2 Ordinal Logit Regression Model

The ordinal logit regression model as shown in equation 3 was utilized to construct a safety prediction model and a comfort prediction model. This model, also known as the proportional odds model is commonly utilized for ordinal dependent variables and was utilized within Yamanaka's study [24]. The maximum number of iterations for the analysis was set at 200. It should be noted that the best results for the model were obtained when PTYPE was treated as a factor within the analysis and not as a covariate variable. Table 9 and Table 10 outlines the results of the regression analysis including the goodness-of-fit and Pseudo R-Square values.

In order to determine the predicted outcomes the equations developed in this study have been shown in equations (4) to (9) below:

$$Prob(k \leq \mathrm{K}) = 1/\{1 + \exp(-\beta_o + \sum_{i=1}^{n} \beta_i X_i)\} \quad (4)$$

$$Prob(k = 1) = Prob(k \leq 1) \quad (5)$$

$$Prob(k = 2) = Prob(k \leq 2) - Prob(k \leq 1) \quad (6)$$

$$Prob(k = 3) = Prob(k \leq 3) - Prob(k \leq 2) \quad (7)$$

$$Prob(k = 4) = Prob(k \leq 4) - Prob(k \leq 3) \quad (8)$$

$$Prob(k = 5) = 1 - Prob(k \leq 4) \quad (9)$$

Where K is the safety or comfort score ranging from 1 to 5, β_o is the parameter estimate value corresponding with each individual threshold for the dependent variable (safety and comfort) as shown in Table 9 and Table 10, β_i is the individual parameter estimate for each independent variable (labelled as location within Table 9 and Table 10), and X_i is the corresponding variable value measured from the field.

The significance values for goodness-of-fit for both the Pearson Rank Correlation, and Deviance tests for both models show values greater than 0.05. This indicates that the model fitting using these measures could be considered, with 95% confidence level, successful. Positive β_i estimates (location) values indicate a positive relationship to comfort or safety, the greater its absolute value the greater the its effect. For instance, within Table 9 under the safety model, the estimate of β_i for LN (LN = MAJ_LN + MIN_LN) is 1.89, indicating that a unit increase in this variable corresponds to an increase in the predicted safety observed by a participant. On the other hand a negative β_i such as -1.74 for NOIS indicates that louder environments have a negative impact on the predicted perceived safety of a participant. The absolute value of β_i for NOIS is less than that for LN indicating that per unit increase in each variable, NOIS has less of an effect on the predicted perceived safety. The odds ratio describes the magnitude of the effect of each variable, an increase of 1 in VOL is associated with a 0.11 decrease in the ordered log odds of being in a lower level of perceived safety. Taking the exponential of the parameter estimates, this reveals that 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.

Some general trends were observed from the model development. First, increased separation from vehicles had a significant positive effect on the rating provided for comfort and safety. Also, less traffic was positively correlated with increase comfort ratings. As expected, the absence of vehicles within the proximity of the participant resulted in higher safety ratings suggesting that the presences of vehicles are the primary concern for cyclists' perception of safety and comfort. Comfort and safety were also observed to be closely correlated suggesting that they share many of the same independent variables. Participants have commented the sense of danger prevented them from feeling comfortable.

Table.**9** Results of Ordinal Logit Regression Model Part 1 (27 Data Points)

Safety Model					**Comfort Model**				
Variables	**Estimate**	**Odds Ratio**	**Wald**	**Sig.**	**Variables**	**Estimate**	**Odds Ratio**	**Wald**	**Sig.**
Threshold									
[SAFTY = 2.00]	-4.528		10.312	.001	**[CMFRT = 2.00]**	-0.93		0.12	0.73
[SAFTY = 3.00]	-3.010		6.594	.010	**[CMFRT = 3.00]**	1.91		0.63	0.43
[SAFTY = 4.00]	-1.702		2.575	.109	**[CMFRT = 4.00]**	3.67		2.13	0.14
Location									
C_SPD_SDEV	-.372	0.69	2.353	.125	**C_SPD_INS**	0.29	1.33	1.60	0.21
					M_SPD	0.02	1.03	0.02	0.88
VOL	-.122	0.89	2.198	.138	**LN**	0.61	1.84	5.07	0.02
					VOL	-0.01	0.99	0.01	0.90
					NOIS	-1.97	0.14	4.57	0.03
Goodness-of-Fit			**Pseudo R-Square**		**Goodness-of-Fit**			**Pseudo R-Square**	
Test	**Chi-Square**	**Sig.**	**Cox and Snell**	.154	**Test**	**Chi-Square**	**Sig.**	**Cox and Snell**	0.38
Pearson	82.004	.299	**Nagelkerke**	.170	**Pearson**	60.29	0.86	**Nagelkerke**	0.42
Deviance	60.647	.901	**McFadden**	.069	**Deviance**	48.10	0.99	**McFadden**	0.20

4.3 Model Validation

Finally, the last step of the analysis was to verify the models developed in part 2 against the remaining 24 data points collected during the IPB runs but not used in model development. Table 11 summarizes the results. The model developed using 63 data points was on average within 0.23 of the actual rating for safety and within 0.95 of the actual rating for comfort. These are encouraging results, especially for a model produced from so few data points. However, these results also suggest that outliers exist in these cases and that a participants feeling of comfort is more subjective than safety. Outliers can be seen by the higher errors in Table 11, and may indicate several notions: first, that there are other independent variables (e.g. intersections or driveways) or confounding factors (e.g. speed) not accounted for in the models. Second, they may suggest that additional data is needed to refine model parameter estimates. Finally, these outliers suggest significant variation between individuals, as there are on most transportation related models (e.g. Highway capacity, 85 percentile speed profiles).

Therefore, careful interpretation of results and statistical validation is required going forward in future research.

Table.10 Results of Ordinal Logit Regression Model Part 2 (87 Data Points – 75% of data)

Safety Model					**Comfort Model**				
Variables	**Estimate**	**Odds Ratio**	**Wald**	**Sig.**	**Variables**	**Estimate**	**Odds Ratio**	**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
NOIS	-1.74	0.18	8.00	0.00	**NOIS**	-2.18	0.11	11.67	0.00
C_SPD_V	0.03	1.03	0.24	0.63	**C_SPD_V**	0.17	1.18	2.33	0.13
S_EXP	0.31	1.37	1.77	0.18	**FIT**	0.78	2.18	1.10	0.29
[PTYPE=1.00]	8.01		1.85	0.17	**[PTYPE=1.00]**	-0.78		0.05	0.83
[PTYPE=2.00]	4.25		2.27	0.13	**[PTYPE=2.00]**	-0.56		0.07	0.79
[PTYPE=3.00]	0				**[PTYPE=3.00]**	0			
Goodness-of-Fit			**Pseudo R-Square**		**Goodness-of-Fit**			**Pseudo R-Square**	
Test	**Chi-Square**	**Sig.**	**Cox and Snell**	.584	**Test**	**Chi-Square**	**Sig.**	**Cox and Snell**	.511
Pearson	174.76	1.00	**Nagelkerke**	0.62	**Pearson**	160.37	1.00	**Nagelkerke**	0.55
Deviance	125.35	1.00	**McFadden**	0.31	**Deviance**	114.24	1.00	**McFadden**	0.28

Table 11. Comfort Model and Safety Model Prediction Validations

Test Sample	Actual Safety	Model Prediction	Error	Actual Comfort	Model Prediction	Error
1	3	3.8	0.8	5	1.4	-3.6
2	5	4.2	-0.8	4	1.4	-2.6
3	3	4	1	4	1.9	-2.1
4	4	4.7	0.7	4	4.7	0.7
5	5	4.5	-0.5	4	2.5	-1.5
6	2	3.4	1.4	3	1.5	-1.5
7	2	3.1	1.1	3	1.4	-1.6
8	4	4.8	0.8	4	4.5	0.5
9	5	4.8	-0.2	4	4.8	0.8
10	5	4.8	-0.2	3	4.4	1.4
11	5	4.8	-0.2	5	4.6	-0.4
12	3	1.7	-1.3	3	1	-2
13	3	1.6	-1.4	4	1	-3
14	5	4.4	-0.6	4	2.5	-1.5
15	5	3.4	-1.6	3	1.5	-1.5
16	5	4.9	-0.1	5	4.9	-0.1
17	4	3.6	-0.4	3	3	0
18	5	3.6	-1.4	5	2.9	-2.1
19	4	3.4	-0.6	5	3.8	-1.2
20	4	3.5	-0.5	5	2.9	-2.1
21	5	3.4	-1.6	5	3.8	-1.2
22	5	5	0	5	4.9	-0.1
23	5	5	0	4	5	1
24	5	5	0	4	4.9	0.9
Average	4.21	3.98	-0.23	4.08	3.13	-0.95

5 CHALLENGES & LIMITATIONS

Although there is inherent risk in offering preliminary results for publication from modelling with incomplete, and small or limited data sets, the value from their disclosure was felt warranted. Moreover, all results should be viewed in context of inherent limitations and challenges faced in this initial test phase.

First, the participant pool is limited to research assistants working in the STS laboratory. Therefore, their individual hypothesis on the results may be reflected in the ratings they provide. Second, assumptions were made in data collection. For instance, traffic volume was not considered for participants cycling on a separated path. The assumption was that on a separated path, the presence of traffic will have a negligible effect on safety and comfort. Third, the maneuverability of the IPB is a concern as it can affect comfort and safety ratings. It was noted that the forward strut extension on which the Kinect sensor and webcam is mounted, made the IPB less maneuverable than a typical bicycle. Modifications to the

IPB for a more ergonomic mounting strut will be made for the next phase of research. Fourth, during the test runs, participants were followed by an investigator for safety and to assist with any issues with the IPB should it arise. It was suspected that safety and comfort ratings provided by the participants may have been slightly higher due to the knowledge that the investigator was cycling with them. To minimize this source of error, future tests will specify a minimum following distance of at least 20 m. Fifth, during test riding, technical challenges with software operation and hardware resulted in discontinuous and missing data in many samples. Continued development of the IPB and its associated software systems will ensure that these issues are resolved for the second phase of testing. Finally, additional variables and/or adjustments to the way current variables are measured will improve the models developed. Additional variables may include driveway visibility, presence of turning vehicles at intersections and presence of large vehicles. Existing variables such as AGE, WIND, NOIS, TEMP, R_EXP and CAR_SPD can be adjusted from an ordinal to a scale variable for more accurate modelling.

SUMMARY & CONCLUSIONS

The intent of the preliminary study on which this paper was based was to prove a proof-of concept regarding IPB data collection and model development methodology. To that end, it would be considered a modest success despite not having full sensory data. These sensors are continually being developed and improved in preparation for UBC's STS Phase 2 data collection program. This study and model development was carried out despite the challenges common to early research programs.

In order to develop BCSPMs a series of test rides were performed on local roads separated into segments of consistent physical characteristics. Data from the test rides was utilized in the development of preliminary models. The IPB developed in this study shows potential for real-time data collection. To fully utilize all the sensors, additional software development and testing is required. In order to successfully develop the model, a substantial amount of data from the general public is desired. More data collected from a wider demographic as well as refinement of the independent variables are required for more accurate models to be developed. After model development, further testing in other geographical locations is possible.

Initial analysis yielded observations comparable to research studied in the literature review. Similar to Yamanaka, Joo and Yao's study [19, 23, 24] the amount of open space in front of the cyclist (F_OBS and LN), the path type (P_TYPE) and the cycling speed (C_SPD_V, T_SPD_V, C_SPD_INS, T_SPD_INS, M_SPD, T_SPD_SDEV, and C_SPD_SDEV) were found to be significant factors for a cyclist's sense of safety and comfort. Other significant variables revealed in this study include traffic volumes and speeds (VOL and CAR_SPD) as well as cyclists experience and fitness (R_EXP, S_EXP and FIT). Phase 2 of the study will focus on the most significant variables and improvement on the models developed herein.

REFERENCES

[1] T. J. Ayres, "Bicycle promotion as a response to climate change", Ergonomics in Design, v 22, p 30-32, 2014.

[2] Intergovernmental Panel on Climate Change, "Mitigation of climate change - Climate change report 2007: Working group III". http:// www.ipcc.ch/publications_and_data/ar4/wg3/en/ch5s5-3-1-5.html. Accessed Sep 1, 2014.

[3] J.J. de Hartog, H. Boogaard, H. Nijland and G. Hoek, "Do the health benefits of cycling outweigh the risks?", Environmental Health Perspective, 118 (8), pp. 1109-1116, 2010.

[4] L. Arundell, "The cost of cycling", Thinking on Two Wheels Cycling Conf, U. of South Australia, 2007.

[5] International Energy Agency, "CO2 emissions from fuel combustion – 2012 Edition", International

Energy Agency Report, Paris, 2012.
[6] U.S. Environmental Protection Agency , “Greenhouse gas emissions from US transportation sector, 1990-2003”, U.S. Environmental Protection Agency – Climate Change, March, 2006.
[7] K. Saelensminde, “Cost benefit analyses of walking and cycling track networks taking into account insecurity, health effects and external cost of motorized traffic”, Transportation Research Part A, 38: 593-606, 2004.
[8] P. Frank, “Chicago regional household travel inventory: Mode choice and trip purpose for the 2008 and 1990 surveys”, Chicago Metropolitan Agency for Planning, June, 2010.
[9] Victoria Department of Transport, “Victorian integrated survey of travel and activity 2007 (VISTA 07)”, Victoria Department of Transport - Research and Data, 2009.
[10]K. Teschke, M.A. Harris, C.C.O. Reynolds, H. Shen, P.A. Cripton and M. Winters, “Exposure-based traffic crash injury rates by mode of travel in British Columbia”, Canadian Journal of Public Health 2013:104(1), pp 75-79, 2013.
[11]SWOV, Victims per Seat per Billion km travelled, Dutch Traffic Safety Research Institute (SWOV), The Hague, Netherlands. swov.nl. Accessed August 7, 2014.
[12]N.Y. Tilahun, D. M. Levinson and K. J. Krizek, “Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey”, Transportation Research Part A, 41: 287-301, 2007.
[13]D. L. Harkey, D. W. Reinfurt and M. Knuiman, “Development of the bicycle compatibility index”, Transportation Research Record 1636, Transportation Research Board, Washington, DC, 1998.
[14]S. Turner, S. C. Shafer and W. P. Stewart, “Bicycle suitability criteria for state roadways in Texas", College Station, TX: Texas Transportation Institute, 1997.
[15]L. Leden, P. Garder and U. Pulkkinen, “Expert judgment model applied to estimating the safety effect of a bicycle facility”, Accident Analysis and Prevention, 32(4), pp 589-599, 2000.
[16]C. Allen-Munley, J. Daniel and S. Dhar, “Logistic model for rating urban bicycle route safety”, (1878) pp. 107-115, 2004.
[17]T. A. Petritsch, B. W. Landis, H. F. Huang and S. Challa, “Sidepath safety model bicycle sidepath design factors affecting crash rates”, Pedestrians and Bicycles, (1982) pp. 194-201, 2006.
[18]W. Du, D. Zhang and X. Zhao, “Dynamic modelling and simulation of electric bicycle ride comfort”, 2009 IEEE International Conference on Mechatronics and Automation, ICMA 2009, August 9, 2009 - August 12, pp 4339-4343, 2009.
[19]M. Yao, Q. Fu, L. Gao and Z. Li, “Exploring influencing factors to the riding comfort of bicyclists on physically separated bicycle roadways in China using proportional odds model”, 11th International Conference of Chinese Transportation Professionals: Towards Sustainable Transportation Systems, ICCTP 2011, August 14, 2011 - August 17, pp 718-727, 2011.
[20]J. Vanwalleghem, I. De Baere, M. Loccufier and W. Van Paepegem, “Sensor design for outdoor racing bicycle field testing for human vibration comfort evaluation”, Measurement Science and Technology, 24(9), 2013.
[21]D.A.M. Twisk, M. J. Boele, W. P. Vlakveld, M. Christoph, R. K. Sikkema, R. Remij and A. L. Schwab, “Preliminary results from a field experiment on e-bike safety: Speed choice and mental workload for middle-aged and elderly cyclists”, Proceedings, International Cycling Safety Conference, Helmond, The Netherlands, 2013.
[22]M. Dozza and A. Fernandez, “Understanding bicycle dynamics and cyclist behavior from naturalistic field data (Nov 2012)”, IEEE Transactions on Intelligent Transportation Systems, Vol. 15, No.1, 2014.
[23]S. Joo and C. Oh, “A novel method to monitor bicycle environment”, Transportation Research Part A, 54: pp 1–13, 2013.
[24]H. Yamanaka, P. Xiadong and J. Sanada, “Evaluation models for cyclists’ perception using probe bicycle system”, Proceedings of the Eastern Asia Society for the Transportation Studies, Vol. 9, 2013.