Institut de recherche
Robert-Sauvé en santé
et en sécurité du travail

Evaluation of Test Methods for Determining Footwear Slip Resistance on Ice Surfaces

Chantal Gauvin
Atena Roshan Fekr
Yue Li
Gordon Wong
Wayne Cheng
David Pearsall
Tilak Dutta

STUDIES AND RESEARCH PROJECTS

R-1136-en

MECHANICAL AND PHYSICAL
RISK PREVENTION

OUR RESEARCH
is working for you !

The Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST), established in Québec since 1980, is a scientific research organization well-known for the quality of its work and the expertise of its personnel.

Mission
To contribute, through research, to the prevention of industrial accidents and occupational diseases and to the rehabilitation of affected workers;

To disseminate knowledge and serve as a scientific reference centre and expert;

To provide the laboratory services and expertise required to support the public occupational health and safety network.

Funded by the Commission des normes, de l'équité, de la santé et de la sécurité du travail, the IRSST has a board of directors made up of an equal number of employer and worker representatives.

To find out more
Visit our Web site for complete up-to-date information about the IRSST. All our publications can be downloaded at no charge.
www.irsst.qc.ca

To obtain the latest information on the research carried out or funded by the IRSST, subscribe to our publications:

- *Prévention au travail,* the free magazine published jointly by the IRSST and the CNESST (preventionautravail.com)
- InfoIRSST, the Institute's electronic newsletter

Legal Deposit

Bibliothèque et Archives nationales du Québec, 2021
ISBN 978-2-89797-189-2 (PDF)

© Institut de recherche Robert-Sauvé en santé et en sécurité du travail, 2021

IRSST – Communications, Strategic Watch and Knowledge Mobilization Division
505 De Maisonneuve Blvd. West
Montréal, Québec
H3A 3C2
Phone: 514 288-1551
publications@irsst.qc.ca
www.irsst.qc.ca

Evaluation of Test Methods for Determining Footwear Slip Resistance on Ice Surfaces

Chantal Gauvin[1], Atena Roshan Fekr[2], Yue Li[2], Gordon Wong[2], Wayne Cheng[2], David Pearsall[3], Tilak Dutta[2]

[1] IRSST
[2] The KITE Research Institute - Toronto Rehabilitation Institute, University Health Network
[3] McGill University

STUDIES AND RESEARCH PROJECTS

R-1136-en

Disclaimer

The IRSST makes no guarantee as to the accuracy, reliability or completeness of the information in this document.
Under no circumstances may the IRSST be held liable for any physical or psychological injury or material damage resulting from the use of this information.
Document content is protected by Canadian intellectual property legislation.

A PDF version of this publication is available on the IRSST Web site.

This study was funded by the IRSST.
The conclusions and recommendations are solely those of authors.

PEER REVIEW
In compliance with IRSST policy, the research results published in this document have been peer-reviewed.

ACKNOWLEDGEMENTS

The authors wish to thank Pierre Drouin, mechanical technician and materials characterization specialist at the *Institut de recherche Robert-Sauvé en santé et en sécurité du travail* (IRSST), for his involvement to the development of the mechanical test method and to the laboratory measurements. Similarly, they wish to express their gratitude to Benjamin Leaker, Kayla Morrone and Jin Li, students at the University of Toronto, for their contribution in data collection at the KITE Research Institute of the Toronto Rehabilitation Institute (KITE).

The authors also wish to thank Jérôme Boutin, scientific professional at the IRSST, for the development of the ice surface temperature measuring equipment (installation of thermistors, development of the data acquisition software). They are also grateful to Nicolas Bastien, scientific professional at the IRSST, and Marie-Christine Stafford, statistician at SolutionStat, Montreal, for their expertise and assistance with the statistical analysis.

The authors extend their thanks to the members of the follow-up committee for their invaluable cooperation and their enthusiasm for this research project. They helped in particular with the selection of the protective footwear used in the study. The authors are also grateful to François Ouellet and Édith Vinet, knowledge transfer advisors at the IRSST, for the support they provided to members of the follow-up committee and the research team.

In addition to an IRSST grant, the work done by KITE was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC, grant numbers RGPIN-2017-06655, DGDND-2017-00097) and the United States National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR, grant number 90RE5005-01-00). NIDILRR is a center within the Administration for Community Living (ACL), in the U.S. Department of Health and Human Services (HHS). The content of this report does not necessarily represent the policy of NIDILRR, ACL, HHS, and endorsement by the Canadian Federal Government should not be assumed.

ABSTRACT

Slipping on ice is one of the main risks of outdoor occupational activities during winter. Ice and freezing rain were involved in 14% of slip, trip and fall (STF) accidents from 2014 to 2016 in Quebec. Many workers who perform outdoor activities rely on their boots to prevent them from slipping. However, choosing the best slip resistant footwear is challenging. Currently, there is no standard test method for evaluating slip resistance of footwear on ice surfaces. The SATRA STM 603 whole shoe tester is used in standard test methods (ASTM International, 2019b, F2913-19; International Organization for Standardization [ISO], 2019, 13287:2019) to evaluate the coefficient of friction (COF) of different shoes on other types of surface (e.g. wet and dry quarry tiles). This apparatus can be used in conjunction with a refrigerated ice tray to evaluate the COF of footwear on ice surfaces. The accompanying SATRA TM144:2011 is a proprietary mechanical method that provides rough guidelines for testing footwear on ice surfaces (frosted ice and smooth ice) with the SATRA STM 603. However, little information has been published about the validity of this method. Alternately, a human-centred method, called the Maximum Achievable Angle (MAA) test, was recently developed using the KITE Research Institute's WinterLab, located at the Toronto Rehabilitation Institute – University Health Network. This method evaluates footwear slip resistance on ice surfaces by measuring the maximum slope participants can walk up and down without slipping.

This study was separated into three phases having the following objectives: **phase 1A** to refine the existing mechanical method by determining ice conditions using the SATRA ice tray, and **phase 1B** to evaluate the repeatability and reproducibility of the results obtained with this method using the SATRA STM 603 whole shoe tester on ice surfaces at two different laboratories: one at the *Institut de recherche Robert-Sauvé en santé et en sécurité du travail* (IRSST) and one at KITE; **phase 2**, to compare the mechanical method with the MAA method for evaluating footwear performance on ice surfaces; and **phase 3**, in the cases of inconsistencies between the two methods, to investigate which method is more reliable for ranking footwear by using another human-centred method. The two ice surfaces used in this study were based on the WinterLab's ice surfaces. The WinterLab's dry ice was a smooth cold ice formed and kept at -5.0 ± 1.0°C with ambient air temperature at 2.5 ± 2.0°C. The WinterLab's wet ice was smooth melting ice formed and kept at -1.5 ± 1.0°C with ambient air temperature at 8.0 ± 2.0°C.

For phase 1A, an ice preparation protocol was developed for the SATRA ice tray and a test protocol was defined based on existing standards. The monitoring of ice tray's ice temperatures using thermistors revealed that the ice surface temperature fluctuated as a function of the refrigeration cycle of the ice tray. These fluctuations showed slightly different patterns between the IRSST and KITE labs. Thus, specific temperature set points and restricted temperature ranges for testing on dry and wet ice surfaces were determined for each lab to ensure that ice temperatures measured by the thermistors were as similar as possible in the two labs (within -6.0 to -5.0°C for dry ice, and within -2.0 to -0.5°C for wet ice) and closest to the KITE WinterLab's ice temperatures. Any frost that formed naturally on the ice surface in ambient conditions was removed by wiping the ice surface with a wet cloth at the beginning of the tests. This helped ensure the SATRA ice surfaces better resemble the smooth surfaces of the WinterLab. For phase 1B, ten types of occupational footwear were tested at both labs, on dry and wet ice surfaces and in different slip modes.

The results from the two labs for boots tested on wet ice were equivalent, both in terms of COF values and footwear ranking based on Bland-Altman analyses (Bland & Altman, 2010). For dry ice, although the footwear ranking was equivalent between the two labs, the COF values obtained at IRSST were systematically higher (by around 0.06) than those obtained at KITE. Limitations of this phase included an inability to control the temperature and the relative humidity in the two labs, which may have impeded the reproducibility of the mechanical method.

In phase 2, each type of women's footwear was tested by four female participants and each type of men's footwear was tested by four male participants using the MAA method. The participants were asked to walk up and down slopes at their own pace on a 4 m walkway in the WinterLab, at KITE, while wearing a safety harness. The surface slope was increased systematically until the participants were no longer able to walk without slipping. Four MAA scores were recorded for each footwear model, each ice surface condition (dry or wet ice) and each direction (descending or ascending) defining the maximum slope the participant was able to walk up or down without slipping. The COF values and footwear rankings obtained on wet ice using the mechanical method were close to those obtained using the MAA method. However, for dry ice, the mechanical method gave a different footwear ranking compared with the MAA method. These observed differences may have been due to differences in ambient conditions that were out of our control (ambient air temperature and relative humidity). Efforts were made to maintain ice surface temperature for the mechanical tests as close as possible to the WinterLab's ice surface temperature. The observed differences between the two methods may also be due to the mechanical method's inability to simulate human gait. Hence, further research is needed to refine the mechanical method for estimating slip resistance performance, in order to improve agreement with human-centred approaches.

In phase 3, two out of the ten types of occupational footwear were selected to be tested on wet ice in the WinterLab, with another human-centred method. A level walking test, developed by the KITE research team and using a passive motion tracking system to detect heel contact and toe-off from the velocity signal, measured the number of times each of five participants slipped while wearing a particular footwear model. On wet ice, the selected boots showed similar slip resistance when tested with the mechanical method (Phase 1B), while their slip-resistance qualities were significantly different when tested with the MAA method (Phase 2). The results of the level walking test, which consist in the number of slips encountered by the five participants during the test, were consistent with the results from the MAA method, and disagreed with the mechanical method.

This study provided a better understanding of the use and limitations of the SATRA ice tray for measuring slip resistance. The results showed that our alternative mechanical method must be further refined to make its results more comparable to human-centred methods. Recommendations have been made to address this issue. This study also demonstrated that conducting tests on different ice surfaces, such as dry and wet surfaces, can be useful as a way of getting a more accurate picture of a boot performance, with both mechanical and human-centred methods.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABREVIATIONS

ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
B&A	Bland-Altman statistical analysis
c	Condition
CI	Confidence interval
CNESST	Commission des normes, de l'équité, de la santé et de la sécurité du travail
COF	Coefficient of friction (dimensionless)
$\overline{COF_f}$	Footwear average COF measured with the mechanical method, all repetitions, operators and slip modes combined
$\overline{COF_{f,m}}$	Footwear average COF of each slip mode measured with the mechanical method, all repetitions and operators combined
CV	Coefficient of variation (%)
EN	European Standard
f	Footwear item
F1 to F10	Footwear tested (10 models)
η_p^2	Partial eta squared
i	Number of a test result
IRSST	Institut de recherche Robert-Sauvé en santé et en sécurité du travail; also refers to one of the two laboratories where mechanical tests were conducted with the SATRA STM 603 whole shoe tester
IRSST_COF	$\overline{COF_f}$ measured at the IRSST lab
ISO	International Organization for Standardization
KITE	KITE Research Institute - Toronto Rehabilitation Institute; University Health Network; also refers to one of the two laboratories where mechanical tests were conducted with the SATRA STM 603 whole shoe tester
KITE_COF	$\overline{COF_f}$ measured at the KITE lab

LoA	Limits of agreement
m	Mode of slipping, or slip mode (flat, forepart or heel)
MAA	Maximum achievable angle
MAA_COF	Equivalent COF calculated from MAA
$\overline{MAA\_COF}_f$	Average footwear MAA_COF measured with the MAA method, all participants and slope directions combined
$\overline{MAA\_COF}_{f,s}$	Average footwear MAA_COF of each slope direction measured with the MAA method, all participants combined
n	Number of test results (observations)
o	Operator
R	Coefficient of correlation
SD	Standard deviation
STF	Slip, trip and fall

1. INTRODUCTION

In Quebec, slips, trips and falls on the same level (STF accidents) top the list of the most frequent types of occupational accidents. For the period 2014–2016, they accounted for 13.1% of all injuries accepted by Quebec's Commission des normes, de l'équité, de la santé et de la sécurité du travail (CNESST), representing compensation payments of $140 million per year (Boucher, 2019). These accidents affect workers across many industries, including truck drivers, material handlers, nurses, cleaners, cooks, teachers and public protection officers (police officers, firefighters, correctional service officers and security guards).

STF accidents generally encompass all accidents related to stumbles, slips, trips, stepping in holes or any other loss of balance, whether the person ends up falling or not. Slips are involved in 40 to 60% of injuries in connection with STF accidents (Courtney, Sorock, Manning, Collins, & Holbein-Jenny, 2001; Gauvin et al., 2015; Manning, Ayers, Jones, Bruce, & Cohen, 1988).

Icy surfaces are one of the main hazards for occupational activities in winter. In Quebec, ice and freezing rain were the most frequent causes of STF accidents for the period 2014–2016, ranking first (14%) just ahead of ground surfaces (10%) and body movement or posture (8%) (Boucher, 2019). During the same period, 46% of STF accidents occurred in winter (December to March). Among workers who have to work outdoor, these percentages may even be higher. This is revealed by the results of a study based on accident data for police officers and crossing guards for the period 2009–2011 (Gauvin et al., 2015). It shows that 80% of slips occurred in winter and that ice- and snow-covered surfaces were involved in over three-quarters of all slips.

Occupational safety and health administrations require employers to eliminate or reduce employee exposure to hazards through engineering, administrative controls, or using personal protective equipment such as slip-resistant footwear. Of course, one of the most effective ways to reduce the risk of slipping on icy surfaces is to spread abrasive. In jobs where this kind of measure isn't really possible, however, workers exposed to winter conditions rely chiefly on their work boots to prevent themselves from slipping (Gauvin et al., 2015; Bagheri, Patel, Li, Morrone, et al., 2019). Footwear slip resistance plays a major role in risk reduction (Di Pilla, 2010; Ells, 2014; Grönqvist, Abeysekera, et al., 2001; Swedler et al., 2015). However, there are very few available methods for assessing slip resistance on icy surfaces. Workers need access to accurate information on footwear slip resistance in winter conditions so they can make informed choices (Bentley & Haslam, 1998; Di Pilla, 2010).

2. BACKGROUND

2.1 Test methods for determining slip resistance of footwear on ice surfaces

A shoe's slip resistance can be assessed quantitatively by measuring the coefficient of friction (COF) between the shoe outsole and a test surface (Strandberg, 1985; Tisserand 1985). The COF is the ratio of frictional and normal forces at the sole-floor interface. Although static friction is expected to be important to prevent the initiation of a slip, and dynamic friction determines whether or not a person can recover their balance after starting to slip (Grönqvist, Abeysekera, et al., 2001), it is generally recognized that dynamic COF is more adequate to evaluate the slip resistance of shoes or floors in contaminated conditions (wet, oily, icy) (Chang, Grönqvist, Leclercq, Myung, et al., 2001; Strandberg & Lanshammar, 1981; Tisserand, 1985).

There are several mechanical devices for measuring the COF. A summary of the different slip testers is given in Di Pilla (2010) and Ells (2014); a critical review of testers and methods is given in Chang, Grönqvist, Leclercq, Myung, et al. (2001) and Chang, Grönqvist, Leclercq, Brungraber, et al (2001). The tribometers are intended primarily for evaluating the performance of different floor surfaces using a pre-prescribed test foot. Whole shoe testers are generally designed to evaluate the performance of different shoes using pre-prescribed surfaces, the most common being quarry tiles and stainless steel covered or not with contaminant, such as water, detergent solution or glycerol (ASTM International, 2019, F2913-19; ISO, 2019, 13287:2019). Whole shoe testers seem to better reproduce the tribological characteristics at the shoe-floor interface compared to several existing tribometers (Chang, Leclercq, Lockhart, & Haslam, 2016).

The STM 603 whole shoe tester, developed by SATRA Technology Centre (Northamptonshire, United Kingdom) (Wilson, 1990; Wilson, 1996), is the more widely used by standard organizations and in footwear industry. The proprietary SATRA TM144 test method, first published in 1992, formed the basis of the European Standard EN 13287:2004, subsequently adopted by the International Organization for Standardization (ISO) as EN ISO 13287:2006, later revised (2012), and currently published as EN ISO 13287:2019. In 2009, ASTM International signed an agreement with SATRA Technology Centre that resulted in the publication of ASTM F2913-11 (Ells, 2014), recently revised as ASTM F2913-19, which is very similar to the proprietary SATRA TM144 test method. In 2009, the Canadian Standards Association (CSA) integrated the EN ISO 13287:2006 for the testing of the slip resistance of footwear in the standard CSA Z195-09. Typically, with the STM 603 device, a test consists in applying a specified normal force to a footwear item placed onto a test surface and then moving the surface horizontally at a constant speed. The frictional and normal forces are measured with load cells. A series of 5 to 10 successive test runs is performed and the average of the last 5 consecutive runs with the variation of less than 10% is considered as the final COF value.

The most recent update of the proprietary SATRA test method (SATRA Technology, 2011) provides rough guidelines to test footwear on ice surfaces. The STM 603 whole shoe tester can be fitted to a refrigerated ice tray (STM 603ICE from SATRA Technology Centre), so that different types of ice surfaces can be created in the laboratory, including frosted ice, dry smooth ice and wet ice. The SATRA TM144:2011 method recommends testing footwear on frosted ice surfaces set at -7°C (the depth of the frost being between 1 and 2 mm), in a controlled lab environment (temperature of 23 ± 2°C; relative humidity of 50 ± 4%). According to the test method, the first and fourth COFs must be reported, representing the COF on frosted ice and on smooth ice (SATRA trainer, personal communication, 2017). The method recommends prior conditioning of the footwear in a cooling bath at -7°C (containing a cooling solution of 50% ethanol and 50% distilled water) for 3 hours. The footwear can also be at a higher temperature than the ice surface, for example 23°C; both variants are valid and may be regarded as complementary. According to SATRA TM144:2011, footwear conditioning can be representative of a step taken very shortly after the wearer leaves a relatively warm environment (for example, leaving a building or a car) or a step taken during a walk in the winter when the soles are colder and may have undergone cold hardening.

However, performing tests with a whole shoe tester on ice surfaces can be challenging, because maintaining specific ice conditions in a laboratory at room temperature for the duration of the tests is difficult, and so is controlling footwear conditioning before beginning a test. Also, little information has been published about this test method using the refrigerated ice tray, and the repeatability and reproducibility of such tests on ice surfaces have not been assessed. An inter-laboratory study performed on COF measurements on contaminated indoor surfaces with this kind of whole shoe tester showed that although a single machine provides an acceptable repeatability limit of 10%, the results of the various machines differ considerably, with a reproducibility limit of up to 70% (Jung & Fischer, 1993). More recently, another inter-laboratory study, carried out on ASTM F2913-19 using STM 603 on dry and wet quarry tiles, revealed improvement in the reproducibility of the results (ASTM International, 2019a). Some studies have evaluated the ability of the STM 603 device to provide a footwear slipperiness measurement representative of the actual footwear performance experienced by human subjects walking on contaminated indoor surfaces (Beschorner, Iraqi, Redfern, Cham, & Li 2019; Blanchette & Powers, 2015; Hunwin, Ormerod, & Darby, 2010).

Human-based approaches have higher validity for assessing footwear function because they take into account the capacity of human beings to adapt their gait to hazardous conditions (Grönqvist, Abeysekera, et al., 2001; Grönqvist, Chang, et al., 2001). A number of evaluations of winter footwear performance have utilized subjective ratings of perceived slipperiness, during which participants or observers rank different types of footwear while participants use them on outdoor winter surfaces (Gao & Abeysekera, 2004; Gard & Berggård, 2006). Some objective measurements have been attempted on natural winter surfaces with mechanical devices such as the stationary step simulator (Gao, Abeysekera, Hirvonen, & Grönqvist, 2004; Grönqvist & Hirvonen, 1995). However, these methods also suffer from inconsistent test conditions as changes in ambient temperature and relative humidity can substantially change the ice properties. Tests of gait and footwear involving stepping or walking by human subjects on slippery slopes have been conducted in previous studies. These studies have greater ecological validity than studies restricted to the use of mechanical devices, but have typically involved only short walkways without snow or ice (Gao, Holmér, & Abeysekera, 2008; Loo-Morrey, 2006). For example, the Health and Safety Laboratories in the United Kingdom have footwear testers walk on a short ramp that progressively tilts to steeper angles until the tester slips (Hunwin, 2010). Unfortunately, this ramp is not designed for testing with ice or snow.

Therefore, new human-centred test methods would be beneficial for assessing winter footwear on icy surfaces and conversely, it remains essential to establish the accuracy of any mechanical methods used. More recently, a human-centred test method has been developed by the KITE Research Institute, at the Toronto Rehabilitation Institute (University Health Network), to evaluate footwear slip resistance on ice surfaces (Bagheri, Patel, Li, Morrone, et al, 2019; Bagheri, Patel, Li, Rizzi, et al., 2019; Hsu, 2015, Hsu, Li, Dutta, & Fernie, 2015; Hsu et al., 2016). These tests are done inside the WinterLab at KITE. This lab has a real ice floor that can reach sub-zero temperatures. The test method, called the Maximum Achievable Angle (MAA) test, measures the steepest incline that participants wearing test footwear can walk up and down without experiencing a two-foot slip (Bagheri, Patel, Li, Morrone, et al, 2019; Bagheri, Patel, Li, Rizzi, et al., 2019). The method can be performed on wet ice or smooth dry ice. The equivalent COF is calculated by taking the tangent of the angle, as in the calculation for a ramp test (James, 1999; Ormerod, 2010).

The MAA test method has demonstrated good repeatability and consistent results when ranking footwear on various surfaces (Hsu et al., 2015; Hsu et al., 2016). This method has also shown to be a useful and valid test in a study where winter boots with a high MAA score worn by workers have actually been less slippery than those of a control group (Bagheri et al., 2019d). Whereas a mechanical test method looks at footwear performance in isolation, the MAA human-centred approach captures users' adaptation of their gait to real ice surfaces. As it is representative of real walking in controlled winter conditions, this method can be regarded as a gold standard for evaluating the accuracy of a mechanical method (Bagheri et al., 2020).

2.2 Classification systems for footwear performance

Measured COF values have been connected to subjective evaluations of slipperiness of different underfoot surfaces (Grönqvist, 1995) or have been associated to slip probabilities (Chang, Matz, & Chang, 2013). However, no generally accepted safe COFs have been established (Di Pilla, 2010; Ells, 2014). Presently, for footwear to be labelled as slip resistant, Canadian standard CSA Z195-14 (R2019) on Protective Footwear (CSA Group, 2019) requires performing tests according to ISO 13287 and only requires that test results be reported on the label. The Canadian standard does not recommend any specific critical threshold. On the other hand, European standards use established codes (SRA, SRB and SRC, see ISO 2012a, ISO/TR 18690:2012) that indicate the shoes have met the specified requirements in standards for safety, occupational and protective footwear (ISO, 2011, 20345-2011; ISO, 2012b, 20347-2012; ISO, 2014, 20346:2014) while tested according to ISO 13287. However, these codes are criticized because they may not always be sufficient for consumers to make an informed choice (Ormerod, 2010).

Using MAA results, KITE developed a rating system for outdoor winter footwear for consumers (www.ratemytreads.com). Footwear that achieves or surpasses the angle of 7°, 11° or 15° receives one, two or three snowflake(s) as a rating unit, respectively (Figure 1).

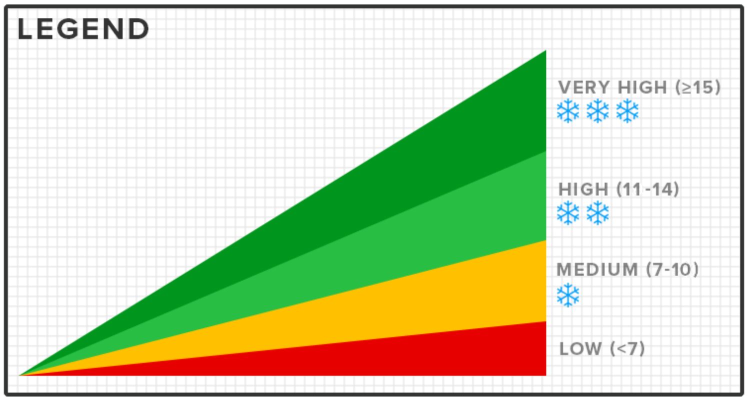


Figure 1. MAA ratings using snowflakes (taken from the web site www.ratemytreads.com with permission).

The 7° cut-off (equivalent to a COF of 0.12) was based on the maximum allowable slope for a curb ramp as specified in existing accessibility guidelines for the built environment (*Accessibility for Ontarians with Disabilities Act,* S. O., 2005, c. 11), with the expectation that footwear should prevent slips on commonly encountered icy curb ramps. The rating system would allow consumers to evaluate potential footwear options in terms of their ability to prevent slips, before purchasing new winter footwear. When the web site was launched in October 2016, around a hundred casual and occupational footwear models were included and only few of them (about 10%) had one snowflake. In 2019, over 200 footwear models are included, and many of them earned one or two snowflakes. Over three years (2016-2019), it was observed that the human-centred MAA slip resistance testing method, with the simple and easy-to-understand snowflake rating system for winter footwear, can be an effective way to provide practical information and promote accident prevention measures based on the sizable local, national and international media interest (such as CBC in 2016[1]). This media coverage resulted in over 100,000 unique visitors to www.ratemytreads.com and increased sales of the footwear with the highest ratings (Summers, personal communication, 2019).

[1] https://www.cbc.ca/news/technology/winter-boots-tested-ice-1.3867531

3. OBJECTIVES

The goal of this study was to evaluate a mechanical method using a whole shoe tester to determine footwear slip resistance on ice surfaces and to compare it with the MAA human-centred test method developed by KITE.

Given that the mechanical method was being compared with the MAA method, which was performed in the WinterLab, the ice surfaces used in the mechanical method had to be as similar as possible to the KITE WinterLab's ice surfaces, for both dry and wet ice surfaces.

The study was separated into three phases having the following objectives:

Phase 1:

A) Develop an alternative mechanical method based on existing test protocols by determining ice conditions closest to the KITE WinterLab ice conditions.

B) Evaluate the repeatability and reproducibility of the novel developed mechanical method for measuring footwear slip resistance at two different laboratories (KITE and IRSST).

Phase 2: Compare results from the mechanical method with the MAA human-centred method for evaluation of footwear slip resistance performance on ice surfaces.

Phase 3: In the cases of inconsistencies between the two methods, investigate which method is more reliable for ranking footwear, by using another human-centred test method with a subset of two types of footwear.

This project also explored how combining the results from different test methods and ice surfaces can lead to recommendations for choosing slip-resistant winter boots.

4. PHASE 1 – EVALUATION OF MECHANICAL TEST METHOD

4.1 Method

4.1.1 Boot selection

Ten models of winter work boots were selected for the study (Table 1). Eight of the models (F1 to F8), chosen with the help of workplace representatives, are used in a variety of workplaces, including police departments, firefighting departments and municipal services. Two other models (F9 and F10), which had already showed promising results in MAA assessments (www.ratemytreads.com), were added shortly after the start of the project, in order to have a full range of boots with different types of performance.

The hardness of the footwear soles was measured (Shore A) using a digital durometer (Shimana, model SHPMDR176, Digital Measurement Metrology, Brampton, ON, Canada). The values reported in Table 1 represent the mean and standard deviation of five measurements taken at five different places in the same polymer material at the sole surface. The specific polymer compounds of the footwear were not known. The boot F9 uses Green Diamond technology which adds metallic particles and other grit embedded into the sole to create a rough surface. The boot F10 uses the Arctic Grip technology that consists in a composite material where glass fibres are incorporated in a rubber material so that these fibres penetrate the ice surface during walking.

For each model of boot, 4 pairs were purchased, either for women in U.S. sizes 7 (2 pairs), 8 (one pair) and 9 (one pair), or for men in U.S. sizes 9 (2 pairs), 10 (one pair), and 11 (one pair). One pair was used for the tests with the mechanical method (left boot only), and three pairs were used for the assessment with the human-centred MAA test method. All the boots were brand new with no prior use.

Table 1. Ten types of winter occupational footwear

Footwear ID	Image	Style	Sole hardness (Shore A)
F1		Men	71.0 ± 1.5
F2		Women	80.5 ± 0.9
F3		Women	71.3 ± 0.9 (black part) 77.2 ± 0.7 (orange part)
F4		Men	73.4 ± 1.1
F5		Men	75.7 ± 0.9
F6		Women	74.9 ± 1.2 (black part)
F7		Women	67.0 ± 0.5
F8		Men	63.4 ± 2.0
F9		Men	64.4 ± 2.3
F10		Men	62.6 ± 0.9 (black part) 73.1 ± 2.0 (Arctic Grip)

4.1.2 Apparatus

The mechanical test method for evaluating the dynamic COFs of footwear on ice surfaces was developed based on the standard test method ASTM F2913-11, which is defined for quarry tiles and stainless steel in dry or contaminated conditions, and the proprietary test method SATRA TM144:2011, which suggests guidelines for testing on ice surfaces.

The apparatus was a whole shoe slip resistance tester (STM 603, SATRA Technology Centre, Northamptonshire, United Kingdom), used in conjunction with a refrigerated ice tray (STM 603ICE, SATRA Technology Centre, Northamptonshire, United Kingdom) (Figure 2). In accordance with ASTM F2913-11, a shoe last (STM603ENL, SATRA Technology Centre, Northamptonshire, UK) of a suitable size for the test footwear was installed in the footwear and then secured in the apparatus. Both IRSST and KITE labs had the same model of SATRA equipment.

The principle of the method consists in applying a vertical contact force of 400 N or 500 N (depending on the footwear size, i.e., 400 N for < 7.5 men's and 8.5 women's US size or 500 N for ≥ 7.5 men's and 8.5 women's US size, as per ASTM F2913-11) to the test footwear against the ice surface. The test surface was then moved horizontally relative to the footwear at 0.3 m/s. For each test run, the values of the vertical (contact) and horizontal (frictional) forces were determined at a time of 0.1 ± 0.01s after the start of the sliding movement. The COF for each test run was calculated by the ratio horizontal force/vertical force. The higher the COF, the better the slip resistance.

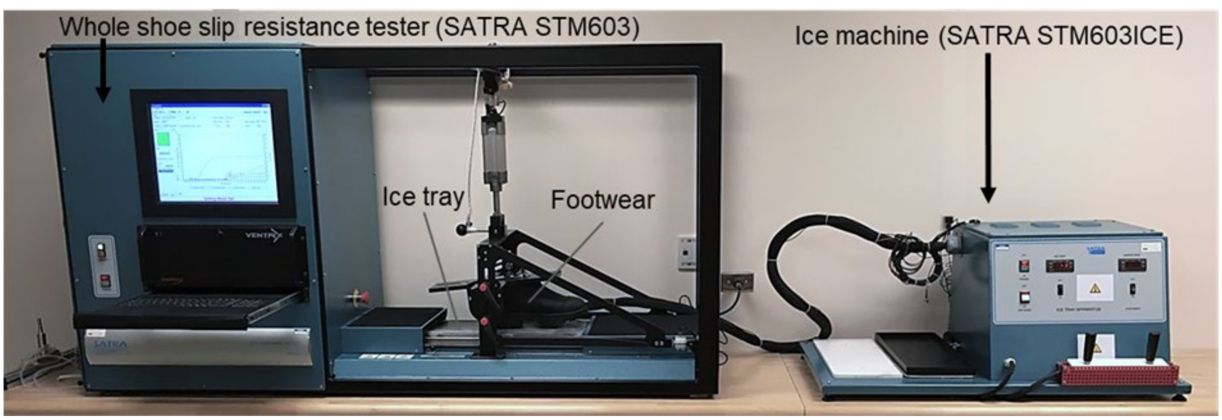


Figure 2. Whole shoe STM 603 test machine and refrigerated ice tray.

With this equipment, the position of the footwear and the line of action of the vertical force with respect to the sole-surface contact area could be set to obtain three test modes: forward heel slip, backward forepart slip and forward flat slip (Figure 3, as per ASTM F2913-11). Though recommended in the SATRA TM144:2011 guideline, the metal wedge was not applied for the heel and forepart test modes, as it tended to melt the ice and sink into it. Instead, the footwear was pressed flat on the test surface, the angle (θ) of the SATRA metal frame was determined, and the metal frame was adjusted to θ + 7° so that the footwear would contact the ice at a 7° angle when the normal force was applied.

(a) Forward heel slip (b) Forward flat slip (c) Backward forepart slip

Figure 3. Three test modes: (a) heel, (b) flat and (c) forward, as described in ASTM F2913-11. The forward heel slip and the backward forepart slip were set with an angle of 7°. The vertical arrow (V) represents the line of action of the vertical force with respect to the sole-surface contact area. The horizontal arrow represents the forward (F) or backward (B) sliding direction of the footwear relative to the surface.

The ice surface was created using the refrigerated ice tray (STM 603ICE, SATRA Technology Centre, Northamptonshire, United Kingdom) (Figure 2). The tray (19 cm x 44 cm x 0.5 cm), containing a cooling coil, was filled with distilled water (400 ml) and the ice machine was set at a specific temperature. The water was then allowed to cool for 1.5 to 3 hours and the ice to form, which continued to remain cool in the ambient laboratory temperature throughout testing time. The machine indicated the ice temperature at all times. From this reading and from the slight noise of the ice machine compressor, a refrigeration cycle could be observed in which the compressor stopped for a while as the ice temperature decreased, then restarted to cool the ice again. This cycle continued around a set point temperature, with a variation range of approximately ± 2°C.

A series of up to 10 successive runs could be performed for each specific shoe in a given experimental condition. The acquisition data software, SlipMASTER, automatically calculated the COF of each test run and the average COF of the last 5 consecutive runs.

4.1.3 Development of the alternative mechanical test method

4.1.3.1 Monitoring ice temperature

To measure the actual ice temperature in the two laboratories, six thermistors (model SC50F103VN, Amphenol Thermometrics, Inc., St. Marys, PA, USA, for the IRSST lab; Mon-a-therm^{TM} temperature probe, Nellcor Puritan Bennett Inc., Pleasanton, CA, USA, for the KITE lab) were placed at different locations: three thermistors underneath the surface of the ice (bottom position), so directly on the ice tray metallic surface, and three others on the ice surface that were set in position by a few drops of water that froze in less than a minute (top position). In-house software programs were developed at both IRSST and KITE labs to provide real-time readings of the thermistors using a data logger (USB-6002, National Instruments, TX, USA, for IRSST lab; Smartreader 8+, ACR Systems, Canada, for the KITE lab).

For two to three cycles, the changes in temperature indicated by the ice machine were noted by hand, then associated with the actual temperatures measured by the thermistors. This temperature monitoring was done after the ice tray had been placed on the metal surface of the STM 603 tester.

An example of the temperatures measured by the six thermistors when the set point temperature of the ice machine was at +2°C at the IRSST and KITE labs is shown in Figure 4.

The figure shows that:

- The temperatures measured on the ice surface were all colder than the set point temperature (which was 2°C in this case) and colder than the temperature reading on the ice machine (which varied from 3°C to 1°C).
- The temperature was not the same at different spots on the surface of the ice (the difference between the bottom and top of the surface was around 1°C; the difference between the lower, middle and upper sections ranged from 1°C to 3°C).
- The highest and lowest temperature zones were not the same for the two ice trays: the warmest zone for the IRSST's ice tray was the upper section, while the warmest zone for the KITE's ice tray was the lower section. Less efficient freezing was observed in these zones.
- The ice surface temperatures varied following a cycle that corresponded to the stop-and-start cycle of the compressor, and that allowed the ice machine to maintain the set point temperature (in this case, the cycle length was approximately 7 min).
- The refrigeration cycles were not the same: at KITE, it seemed that the ice warmed up quickly, but took a while to cool, whereas at the IRSST it was the reverse, with the ice warming up quite slowly, but cooling quickly.

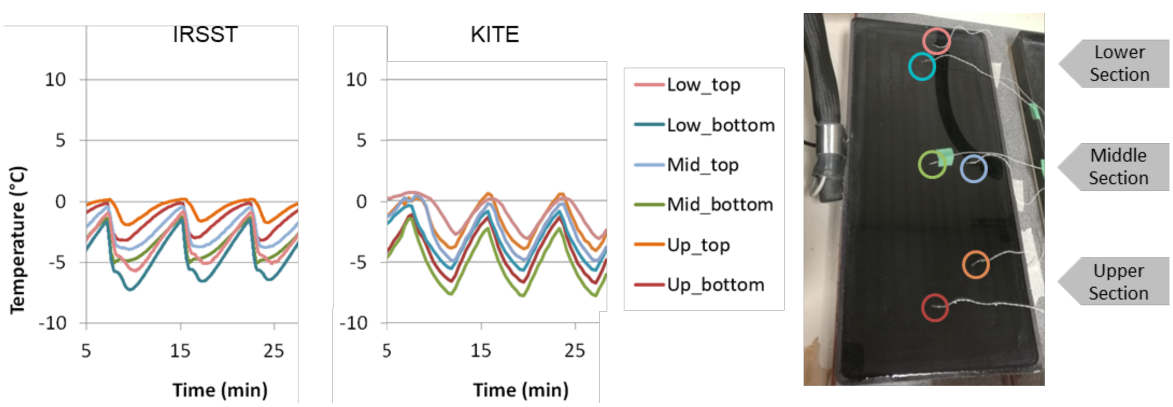


Figure 4. Example of readings of six thermistors positioned underneath (bottom) and on (top) the ice surface, at upper, middle and lower sections of the tray of two ice machines (at IRSST and at KITE), for a set point temperature of 2°C. The Mid_top thermistor (in light blue) is the one used to determine the ice conditions closest to the KITE WinterLab ice conditions.

4.1.3.2 Determination of dry and wet ice conditions

To choose the most appropriate set point temperatures for making wet and dry ice, only the thermistor on the ice surface, at the centre of the test area (middle section), was used (Mid_top, in light blue on Figure 4) and its temperature was compared with the targeted temperatures of the WinterLab:

- Wet ice (WinterLab): melting ice at -1.5°C on average (air temperature: 8°C)
- Dry ice (WinterLab): cold smooth ice at -5.5°C on average (air temperature: 2.5°C).

Different set point temperatures of the ice machine were tried (APPENDIX A). Those with the best potential for producing ice surfaces similar to those of the WinterLab were:

- Wet ice (ice tray): Set point temperature of 2°C at IRSST and 1°C at KITE
- Dry ice (ice tray): Set point temperature of -2°C at both the IRSST and KITE labs.

The ice refrigeration cycles were examined in greater details with respect to these set point temperatures (Figure 5). The sweet spots in the cycle for taking COF measurements are identified by the shaded areas in Figure 5. During these test windows, the mean temperatures at the ice surface were as close as possible to the WinterLab temperatures (difference of less than 1°C for dry ice and less than 1.2°C for wet ice, Table 15, APPENDIX A).

The thermistors were used to determine the offset between the actual ice temperature and the temperature reading on the ice machine. Once the ice conditions had been determined, the ice for the later tests was prepared solely on the basis of the temperature reading of the ice machine, without the use of the thermistors.

Frost formed naturally on the ice surface, especially when the laboratory relative humidity level was high and the ice was exposed for a long time to the ambient temperature. Neither of the two laboratories has a temperature and humidity control system. The measured temperature and relative humidity during testing were respectively about 23°C and 33% at KITE, while they were respectively 20–21°C and 32–58% at IRSST. To make the ice surface more like the smooth surface of the WinterLab, the frost was removed regularly by wiping the surface with a wet cloth (Figure 6). At both labs, a microfibre cloth was used for this purpose.

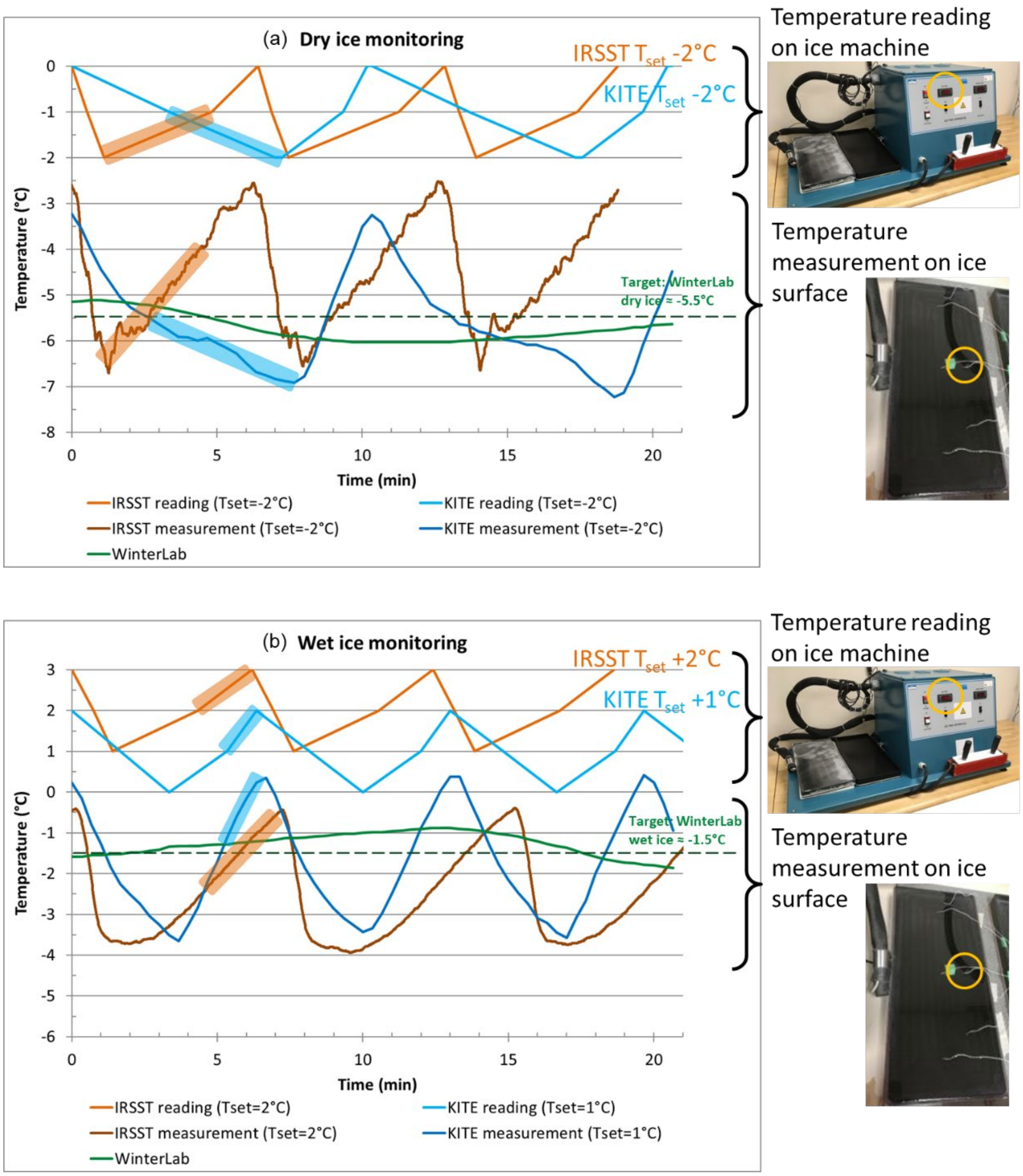


Figure 5. Variation of SATRA ice tray and WinterLab ice surface temperatures for (a) dry ice and (b) wet ice conditions at both labs (IRSST and KITE). In each graph, top curves represent ice tray display temperatures, and bottom curves actual ice surface temperatures (thermistor Mid_top). Temperature variations are in orange for IRSST, in blue for KITE and in green for WinterLab. The ice tray set temperatures are indicated in top-right corners of the graphs. Shaded areas indicate selected test windows.

Figure 6. Wiping the ice surface with a wet cloth to remove the frost.

4.1.3.3 Footwear conditioning

The boots were left at the ambient temperature of the laboratory for testing.

The SATRA TM144:2011 method suggests placing the boot in a bath containing a cooling solution (50% ethanol and 50% distilled water) at -7°C for 3 hours before the testing. The sole of the boot must then be quickly wiped and dried and installed on the test device. Under these conditions, it is not easy to maintain the cold temperature of the boot sole and ensure that it is always the same for all the tests. It turned out that the sole warmed up by approximately 1°C every minute, for the first 10 minutes after it was removed from cold conditioning (APPENDIX A).

The hardness of the soles of the 10 boot models was measured at different temperatures and was found to remain pretty constant between 0 and 23°C. So the hardness of the soles, whether they were cold (as in the case of the WinterLab) or temperate (as in the case of the IRSST and KITE labs), should not have a major impact on the COF measures obtained from the two methods: mechanical and MAA.

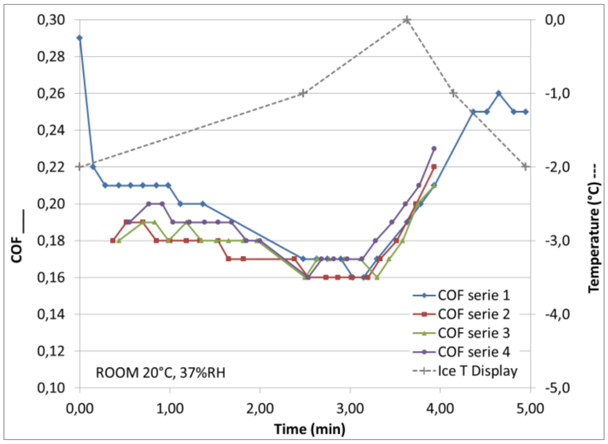
4.1.3.4 Preliminary tests

Before the test protocol was finalized, two series of preliminary tests were conducted, at IRSST only, to see the effect of ice refrigeration cycle on COF measurements, to determine the experimental unit to be considered (mean of last 5 consecutive runs or a specific run as described in SATRA TM144:2011) and to assess the appropriateness of opting for dry ice over frosted ice.

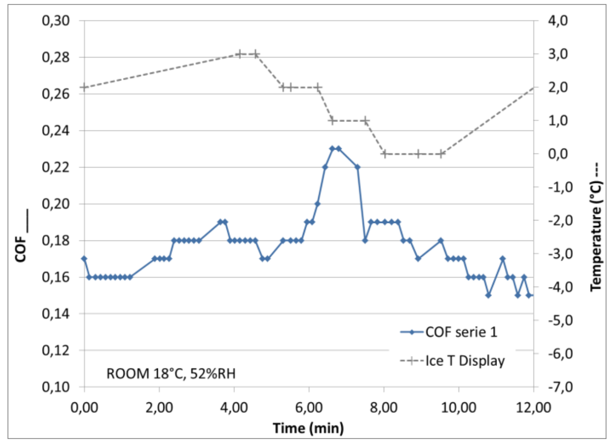
Effect of the refrigeration cycle

The first series of tests involved taking a number of consecutive COF measurements (so disregarding the test windows identified earlier) with a boot (F10 in heel mode) over 4 cycles for dry ice and over 1 cycle for wet ice.

The results show that for dry ice (Figure 7a), the COF declined when the temperature rised, increased when the temperature dropped, and that pattern recurred over the 4 cycles. In addition, deterioration in the boot-ice interface was not observed, as the measured COF values were repeated over the 4 cycles. For wet ice (Figure 7b), the variation in COF was also influenced by the temperature of the ice, but with no clear tendency.



a) Dry ice



b) Wet ice

Figure 7. Variation of ice tray display temperature and COF measured for boot F10, heel mode, in (a) dry and (b) wet ice conditions at the IRSST lab.

COF measurements were fairly sensitive to the temperature of the ice machine, and therefore of the ice surface. But at the very start of the test windows identified earlier (for IRSST, dry: from -2 to -1°C; wet: from 2 to 3°C), the variations in COF were a little less pronounced (the test runs more stable) compared to the other points' variations in the cycle.

Comparison between dry ice and frosted ice

The second series of tests involved taking measurements with two boots (F1 and F10, in heel mode), following the method developed in this study on dry ice and following the SATRA TM144:2011 method on frosted ice. For the two methods, the boots were conditioned at the laboratory ambient temperature, and 3 series of 10 successive runs were conducted with each boot.

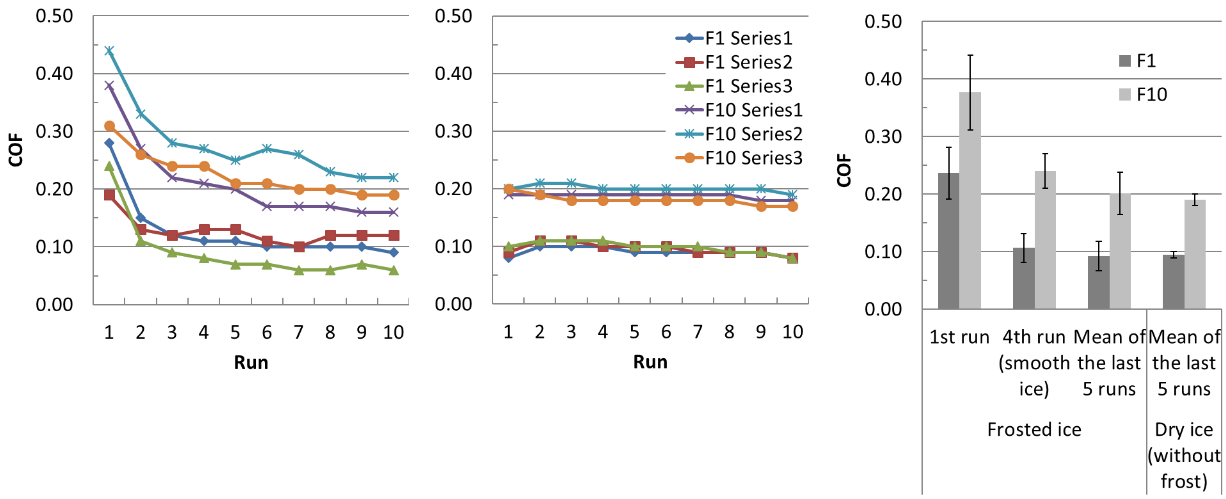
For the SATRA TM144:2011 tests, the temperature for making the ice was set at -2°C, not the suggested -7°C, for the purpose of comparison with dry ice (at -2°C), and in accordance with the recommendations of the ice machine calibration report.[2] The tests were run in accordance with the SATRA TM144:2011 method, on a frosted ice surface, when the depth of the frost was between 1 and 2 mm. The ice was resurfaced between two test series by passing the electrically heated ice dressing tool over the surface of the ice and waiting long enough for the frost to form again. The 1st run corresponded to the COF on frosted ice and the 4th run to the COF on dry smooth ice.

The results (Figure 8) showed that:

- The COFs on the frosted ice (1st run) were higher than those on dry ice (Figure 8c).

[2] The offset observed between temperature set point and actual temperature of the ice was confirmed by the calibration carried out by SATRA Technology Centre on IRSST's ice tray. The calibration report indicated a difference of 5°C between the set point and the temperature of the empty ice tray measured with an infrared reader. So, SATRA Technology Centre suggested using a set point of -2°C to produce ice at -7°C.

- On the frosted ice (Figure 8a), the COFs declined with the runs, until almost reaching a plateau that converged with the COF on dry ice.
- On dry ice (Figure 8b), the COFs seemed to remain fairly stable over the 10 successive runs, with a slight decrease over time (after the 2nd run).
- The repeatability between the three tests was better (i.e. less variation) on the dry ice than on the frosted ice.



a) Frosted ice protocol according to SATRA TM144:2011 b) Dry ice protocol developed in the current study c) Comparison of COF (mean ± SD)

Figure 8. COF evaluated in heel mode for boots F1 and F10 at the IRSST lab on (a) frosted ice condition, (b) wiped dry ice condition (3 series of 10 test runs for each condition) and (c) comparison of mean COF (± SD) evaluated following different protocols.

The decline in the COFs on frosted ice was expected because the boot sole had good traction on the frost on top of the ice. Over time the frost was scraped from the ice, making the ice smoother and more slippery.

The greater variation between the three runs on the frosted surface can be explained by the fact that it was difficult to control the thickness of the frost and to measure it. Furthermore, a frost thickness of 1 mm does not offer the same slip resistance as a frost thickness of, for example, 2 mm.

Determination of experimental unit

The preliminary test results showed that smooth ice without any frost (removed with a wet cloth) could produce repeatable results and the mean of the last 5 consecutive runs could be used as an experimental unit, as is suggested in ASTM F2913-11, contrary to SATRA TM144:2011 which considered the 4th run on frosted ice as the final readout (Figure 8).

However, sometimes some boots showed a greater degree of variation over the course of successive runs. As the software does not automatically calculate the variation of the last 5 runs, the variation was estimated *a posteriori* for the 540 conditions tested in the 2nd part of the phase 1 to see whether it was indeed 10% or less in most cases. The results, presented in APPENDIX A (Table 17), showed that the coefficients of variation were 10% or less for practically all the measurements taken on dry ice (99%), but for only two thirds (67%) of those taken on wet ice. The greater variation in the last 5 runs of the measurements taken on wet ice may be explained by the fact that the measured COFs were lower and very close to the resolution of the data acquisition system of the STM 603, that is, 0.01. For wet ice, the coefficients of variation were more frequently ≤ 10% at the IRSST lab (84% of cases) than at the KITE lab (34% of cases).

4.1.4 Final test protocol

The mechanical test method developed in the first part of phase 1 was used to evaluate the dynamic COF of footwear. The tests were performed with a normal force of 500 N for all men's footwear (US size 9) and 400 N for all women's footwear (US size 7). The ice surface was used to determine the tare weight.

The tests were conducted on dry and wet ice conditions. The ice was wiped with a wet cloth to remove the frost at the beginning of each series of tests, just before the test window, at a time determined by the temperature displayed on the ice machine. Table 2 summarizes the preparation of the ice surfaces prior to the tests (including when to wipe the ice with a wet cloth and when to perform the tests) and the characteristics of the refrigeration cycles at the two laboratories.

It is noteworthy to mention that in some cases, wiping the ice with a wet cloth also served to resurface the ice. After testing boot F9, which has grit embedded into the sole material to create a rough surface, the ice surface had to be wiped several times with a cloth saturated with water in order to remove the scratches made by the sole of the boot (Figure 9).

A series of 5 to 10 successive runs was performed for each specific boot in a given experimental condition. The average of the last 5 consecutive runs, generally showing a variation of less than 10%, was considered as the final COF of the footwear and as the basic experimental unit.

The research team ensured there was as much similarity as possible on the application of the test protocol between the two laboratories.

Table 2. Summary of ice preparation protocol for mechanical test method on dry and wet ice surfaces at IRSST and KITE, established essentially on the basis of ice tray temperatures

	Dry ice		Wet ice	
	IRSST	**KITE**	**IRSST**	**KITE**
Ice making (1.5 hours)				
Ice tray set temperature	**-2°C**	**-2°C**	**2°C**	**1°C**
Ice surface preparation for testing				
Wiping the ice with a wet cloth once before the runs at a specific ice tray temperature	-1°C	0°C	1°C	0°C
Temperature range for testing	**-2°C to -1°C** When the compressor has stopped, and the ice is warming up	**-1°C to -2°C** When the compressor is running, and the ice is cooling down	**2°C to 3°C** When the compressor has stopped, and the ice is warming up	**1°C to 2°C** When the compressor has stopped, and the ice is warming up
Testing time slot	~ 3 min	~ 3 min	~ 2 min	~ 1 min
Characteristics of refrigeration cycle				
Variation of ice tray display temperature	0°C to -2°C	0°C to -2°C	3°C to 1°C	2°C to 0°C
Cycle time	~ 6–7 min	~ 11–12 min	~ 6–7 min	~ 6–7 min
Average ice surface temperature during testing windows (as estimated by thermistors)	~ -5.1°C	~ -6.3°C	~ -1.9°C	~ -0.4°C

Figure 9. Deep scratches made by F9 sole (having the Green Diamond technology) on wet ice.

4.1.5 Testing conditions

The ten types of occupational winter footwear (Table 1) were tested in the three slip modes (flat, forepart, heel) on the two ice surfaces (dry and wet) for a total of 60 testing conditions per lab (IRSST and KITE) (Figure 10). The 60 testing conditions were repeated three times at KITE by one operator, and six times at IRSST by two operators (three times per operator), giving a total of 540 COFs.

For each type of boot, the COF evaluation was done using the same boot specimen, left foot only, for all the tests at both labs. The upper part of each boot was cut off so that only the lower part with the laces was left (Figure 11). Appropriate-size lasts (STM603ENL, SATRA Technology Centre, Northamptonshire, UK) were inserted into the boots to form two subsets:

- Group A: Last sizes 36 (F6), 37 (F7), 39 (F9), 40 (F5) and 41 (F10) were used
- Group B: Last sizes 36 (F2), 37 (F3), 38 (F8), 39 (F4) and 40 (F1) were used.

The tests were carried out in a semi-random order, i.e., for a given ice surface, the footwear models were tested in a random sequence within their group (A or B) for each mode of slipping. They were conducted first at KITE and then at IRSST over several days. Up to 30 tests could be performed in one day (e.g., 5 models of boot x 3 modes x 1 ice surface x 2 operators x 1 repetition).

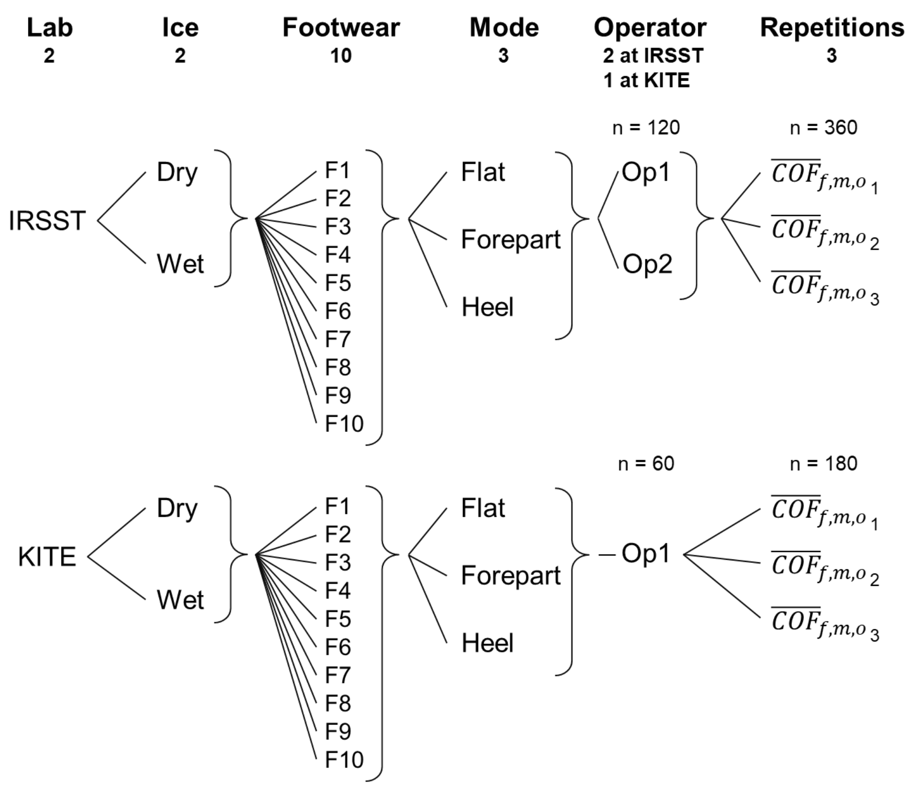


Figure 10. Mechanical testing conditions and frequency.


Figure 11. Boot laces and stuffing used to keep the lasts from shifting inside the boots.

The sole of the boot was abraded with silicon carbide paper wrapped around a rigid block exerting minimal pressure (as per ASTM F2913-11) only once (at the beginning of this study). The sole was washed frequently with detergent solution (as per ASTM F2913-11). No visible wear of the sole was found during the tests because the ice surface is not particularly abrasive; it is rather smooth and slippery. Thus, the integrity of the footwear sole was considered the same for each test condition. The footwear was conditioned at the laboratory temperature for testing.

The laboratory temperature and relative humidity during testing were 16 to 23°C (median of 21°C) and 25 to 58% (median of 42%), respectively, at the IRSST lab, and 23°C and 33%, respectively, at the KITE lab.

4.1.6 Analyses

The analyses were performed separately for each ice surface. According to the results, the COF value in a given condition was established as the mean obtained by grouping together the values (basic experimental units) of the three repeats of the same operator (n = 3), or the repeats of the same lab (n = 6 for IRSST, n = 3 for KITE), or by grouping together the values of all the modes of slipping (n = 18 for IRSST, n = 9 for KITE). The mean, standard deviation (SD) and coefficient of variation (CV) of the COFs were calculated using the following equations:

$$\overline{COF}_c = \frac{1}{n}\sum_{i=1}^{n} COF_{c,i} \quad \textbf{(1)}$$

$$SD_c = \sqrt{\frac{\sum_{i=1}^{n}\left(COF_{c,i} - \overline{COF}_c\right)^2}{(n-1)}} \quad \textbf{(2)}$$

$$CV\,(\%) = \frac{SD_c}{\overline{COF}_c} * 100 \quad \textbf{(3)}$$

where c represents the condition taken into account, i is the number of the test result and n shows the total number of test results, as follows:

- $c = f,m,o$ for each footwear, mode and operator, n = 3 ($\overline{COF}_{f,m,o}$)
- $c = f,m$ for each footwear and mode, n = 6 for IRSST and n = 3 for KITE ($\overline{COF}_{f,m}$)
- $c = f$ for each footwear, n = 18 for IRSST and n = 9 for KITE ($\overline{COF}_{f}$)

where o, m and f represent the operator, the mode (flat, forepart, heel) and the type of footwear (F1 to F10), correspondingly.

To facilitate the interpretation and the analysis of boot performance, the $\overline{COF}_{f}$ obtained were compared with the threshold of 0.12, defined by one snowflake according to the ranking system determined by KITE for the MAA method (threshold angle of 7°). In addition, the boots were divided into two subsets of 5 boots each, according to whether their $\overline{COF}_{f}$ were high (in blue) or low (in red) in dry ice conditions.

The **intra-operator repeatability** of the mechanical method is the variability of the COFs measured by an operator using the same machine. It was assessed by calculating the SD and the CV (equations (2) and (3)) obtained for each of the three operators (two at IRSST and one at KITE) in the 30 conditions (10 boots x 3 modes). The lower the standard deviation (and therefore the coefficient of variation), the greater the repeatability.

To get a better assessment of the overall repeatability for all conditions, the pooled SD (Cohen, 1988) was estimated by equation (4):

$$SD_{pooled} = \sqrt{\frac{SD_1^{\,2} + SD_2^{\,2} + \cdots + SD_k^{\,2}}{k}} \quad \textbf{(4)}$$

where k represents the number of data groups pooled, that is:

- k=10 for SD_{pooled} per operator per mode (for 10 boots)
- k=30 for SD_{pooled} per operator (for 10 boots x 3 modes)
- k=90 for SD_{pooled} per ice surface (for 10 boots x 3 modes x 3 operators)

The **inter-operator reproducibility** of the mechanical method is the variability of the COFs measured by two operators using the same machine. It was established only for the IRSST lab, where two operators took measurements using the same equipment. It was assessed first by calculating the SD and the CV (equations (2) and (3)) obtained for the six repeated measurements of each of the 30 conditions (10 boots x 3 modes), and then by calculating the pooled SD (equation (4)), for each of the ice conditions.

Then an analysis of covariance (ANCOVA) was done to test the effect of the various independent variables (footwear, ice, mode, operator, including the two-way interactions) on the measurement of the COFs, regardless of the effect of temperature and relative humidity in the laboratory, treated as covariates. The purpose of this analysis was to check whether the operator and the laboratory environment (temperature and relative humidity) could have a significant (undesirable) effect on the measurement of COFs.

The **inter-laboratory reproducibility** is the variability of the COFs measured using two machines, i.e., the ones in the IRSST and KITE labs. It was assessed first by different analyses of variance (ANOVA):

- An ANOVA for all the data (n = 540) with the independent variables footwear, ice, mode and lab, as well as the two-way interactions.
- An ANOVA by type of ice surface (n = 270) with the independent variables footwear, mode and lab, as well as the two-way interactions.
- ANOVAs by type of ice surface and by laboratory with the independent variables footwear, mode and operator for the IRSST lab (n = 360), and footwear and mode for the KITE lab (n = 180), as well as certain two-way interactions.

The intent was to identify what the significant effects were and the effect sizes of different independent variables, especially the laboratory effect.

Then, Tukey multiple comparison tests were run by type of ice surface and by laboratory to determine the significant differences between boot COFs, as well as modes of slipping. Performances were ranked by significantly different subsets, and the rankings obtained at the two labs were compared. These tests were done first by taking into consideration each mode of slipping separately (using $\overline{COF_{f,m}}$), and then by combining all the modes of slipping (using $\overline{COF_f}$).

The ANOVA and Tukey statistical analyses were performed using SPSS statistical software version 23 (IBM Corporation, 2015). The analyses had a significance level of 0.05. The assumptions underlying the use of the models (homogeneity and normality of the residuals) were verified by examining the models' standardized residuals.

Last, Bland-Altman (B&A) analyses (Bland & Altman, 2010) were done to gauge the agreement between the measurements taken in the two labs. By plotting the differences between the two series of measurements (IRSST_COF and KITE_COF) in relation to the means of the two measurements, the analyses served to estimate the **bias** between the two series of measurements (i.e., the mean of the differences) and the **limits of agreement** (LoA), where 95% of the differences are found. Thanks to these analyses, it was also possible to estimate the linear correlation (coefficient of correlation R) between the two series of measurements. The measurements taken at IRSST and at KITE were considered to be in agreement if the bias was almost nil (close to zero), the limits of agreement were very low and the coefficient of correlation was close to 1.0.

The B&A analyses were conducted with NCSS statistical analysis software (NCSS 11 Statistical Software, 2016, NCSS, LLC., Kaysville, Utah, USA, https://www.ncss.com/software/ncss/) (Hintze, 2004) done by combining all the modes of slipping using $\overline{COF}_f$ (Design 1 comparison in NCSS 11). They were also done taking each mode of slipping into consideration separately with $\overline{COF}_{f,m}$, taking multiple repetitions of each series of measurements into account (Design 2 comparison in NCSS 11).

4.2 Results

4.2.1 Intra-operator repeatability

For dry ice, SDs varied between 0.000 and 0.040 (CVs between 0 and 29%) depending on conditions, and the pooled SD for all three operators was 0.014 (Table 18, APPENDIX B). For wet ice, SDs varied between 0.000 and 0.020 (CVs between 0 and 47%) depending on conditions, and the pooled SD for all three operators was 0.009 (Table 19, APPENDIX B).

Overall, the intra-operator repeatability for dry ice generally produced CVs of less than 10% (in 73% of cases), or else, at least less than 15% (93% of cases). However, CVs were generally higher for wet ice (CV ≤ 10% in 36% of cases, and ≤ 15% in 64% of cases).

4.2.2 Inter-operator reproducibility (IRSST)

For dry ice, SDs varied between 0.005 and 0.043 (CVs between 2 and 16%) depending on the condition, and the pooled SD was 0.018 (Table 20, APPENDIX B). For wet ice, SDs varied between 0.003 and 0.020 (CVs between 3 and 52%), and the pooled SD was 0.013 (Table 20, APPENDIX B). Overall, inter-operator reproducibility yielded CVs for dry ice that were generally less than 10% (in 70% of cases), or else, at least less than 15% (97% of cases). For wet ice, however, the CVs were higher (CV ≤ 10% in 23% of cases, and ≤ 15% in 40% of cases).

The ANCOVA (Table 3) showed that the covariates *temperature* and *relative humidity* did not have a significant effect ($p > 0.05$) on the measured COFs, within the temperature and relative humidity ranges of the IRSST laboratory during the tests. The analysis also showed that the *operator* did not have a significant effect ($p = 0.767$) on the measurements. This analysis confirmed that the inter-operator reproducibility was acceptable. The two operators at the IRSST lab were therefore not differentiated in the subsequent analyses.

Table 3. ANCOVA for mechanical test results at IRSST including the effect of footwear, mode, ice, and operator, as well as environment as covariate variables (significant with p ≤ 0.05)

Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
Footwear	.766	9	.085	322.482	<.001	.902
Ice	.627	1	.627	2378.118	<.001	.883
Mode	.055	2	.028	104.823	<.001	.400
Operator	2.326E-5	1	2.326E-5	.088	.767	<.001
Footwear * Ice	.291	9	.032	122.397	<.001	.778
Footwear * Mode	.040	18	.002	8.495	<.001	.327
Ice * Mode	.019	2	.009	35.713	<.001	.185
Temperature	2.212E-5	1	2.212E-5	.084	.772	<.001
Relative humidity	4.677E-4	1	4.677E-4	1.773	.184	.006
Error	.083	315	2.637E-4			
Total (corrected)	2.135	359				

4.2.3 Inter-laboratory reproducibility (IRSST and KITE)

The ANOVA done on all the data showed that all the main factors (*footwear*, *ice*, *mode* and *lab*) and the two-way interactions had a significant effect on COF measurements ($p < 0.001$, Table 24, APPENDIX C). To break down the magnitude of the effects, given that everything is significant, the effect size indicated by partial eta squared (η_p^2) was considered[3]. It can be seen that the factors *mode* and *lab* have a lesser effect (η_p^2 = 0.453 and 0.492) than the factors *footwear* and *ice* (η_p^2 = 0.897 and 0.832).

The ANOVAs by type of ice showed that all the factors (*footwear*, *mode* and *lab*) and the two-way interactions again had a significant effect on the measured COFs ($p < 0.001$, Table 25, APPENDIX C). *Footwear* is one of the main factors having the greatest effect on COFs ($\eta_p^2 \approx 0.9$), which is desirable, given that the mechanical method serves specifically to measure, compare and rank the performance of different boots. The *mode* factor also had a significant effect, although its effect size ($\eta_p^2 \approx 0.5$) was less than that of *footwear*. The *lab* factor, albeit significant, still had less of an effect size on wet ice (η_p^2 = 0.069) than on dry ice (η_p^2 = 0.764). This means that the two labs produced different COFs, but these differences were more noticeable for measurements on dry ice than on wet ice.

[3] Partial eta squared is a measurement of effect size commonly used in ANOVAs. Varying between 0 and 1, it refers to the strength of the association between two variables. To interpret this value, Cohen's convention cannot be applied here, as it was developed in a different research context. The value was instead interpreted relatively over all the factors (large effect size ≈ 0.9; medium effect size ≈ 0.5; low effect size ≈ 0.1).

Although the mode of slipping had a significant effect, the effect was not sufficient to substantially change the tendencies observed between the two labs. In other words, the comparison of measures between the two labs was essentially the same, regardless of the mode of slipping, with a few exceptions. Hence, the two following sections present analyses that were done using overall COFs ($\overline{COF_f}$). The analyses by mode of slipping (with $\overline{COF_{f,m}}$) are presented in APPENDIX C, to which reference will be made when relevant.

4.2.3.1 Dry ice

The ANOVAs conducted for each laboratory for dry ice (Table 4) showed once again that the main factors (*footwear* and *mode*) and the interaction between these two factors explain most of the differences between the COFs measured for the two labs. Partial eta squared indicates a large effect size for the main factor *footwear* (η_p^2 > 0.9). For the *mode* factor (and its interaction with the *footwear* factor), the effect size was less for the measurements taken at IRSST (η_p^2 ≈ 0.5) than for those taken at KITE (η_p^2 ≈ 0.8). A significant interaction between the two factors indicates that the COF values estimated for boots in one mode of slipping may be different from those estimated in another mode of slipping.

Table 4. ANOVAs for mechanical test results on dry ice for both labs (IRSST and KITE), including effect of footwear and mode (significant with p ≤ 0.05)

Lab	Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
IRSST	Footwear	.490	9	.054	177.213	<.001	.915
	Mode	.062	2	.031	100.707	<.001	.575
	Operator	2.404E-4	1	2.404E-4	.783	.378	.005
	Footwear * Mode	.043	18	.002	7.715	<.001	.482
	Error	.046	149	3.070E-4			
	Total (corrected)	.640	179				
KITE	Footwear	.220	9	.024	250.605	<.001	.974
	Mode	.033	2	.016	169.324	<.001	.849
	Footwear * Mode	.014	18	.001	7.850	<.001	.702
	Error	.006	60	9.738E-5			
	Total (corrected)	.272	89				

Through Tukey multiple comparisons, significant differences between the boots could be assessed, and are presented in the form of subsets in Table 5. The boots in blue obtained $\overline{COF_f}$ significantly higher than those in red, for both laboratories. The darker line in the Table 5 separates the boots that obtained $\overline{COF_f}$ ≥ 0.12 from others. The $\overline{COF_f}$ values recorded at IRSST were higher than those recorded at KITE for all boots. The results also indicate that the COFs from 8 boots were greater than 0.12 at IRSST, and from 5 boots at KITE.

Table 5. Footwear ranking and significant subsets according to overall COF obtained with mechanical test method on dry ice for both labs (IRSST and KITE)

Foot wear	N	Subsets for **IRSST** – Overall COF				
		1	2	3	4	5
F1	18	0.107				
F2	18	0.119				
F3	18		0.142			
F6	18		0.151			
F5	18		0.158			
F9	18			0.189		
F7	18				0.222	
F4	18				0.224	
F10	18				0.226	
F8	18					0.277

Foot wear	N	Subsets for **KITE** – Overall COF					
		1	2	3	4	5	6
F2	9	0.048					
F1	9		0.076				
F3	9		0.084				
F6	9		0.086				
F5	9			0.104			
F9	9				0.134		
F4	9				0.148	0.148	
F7	9					0.159	
F8	9						0.195
F10	9						0.202

Overall, the ability of the method to discriminate between the boots was equivalent from one lab to the next, or from one mode to another (APPENDIX C), with 5 or 6 subsets. The ranking of the boots was fairly equivalent at the two labs.

The significant differences in the COF values between the modes of slipping, estimated with all boots, are presented in Table 6. The lowest COF values were generally measured with respect to the forepart mode of slipping and the highest values were observed in flat mode in both labs. The three modes were significantly different for the IRSST lab, with differences of approximately 0.02 between the modes. At the KITE lab, modes forepart and heel produced similar COFs, but with a difference of approximately 0.04 compared to the COF in flat mode.

Table 6. Significant differences between modes according to COF (all footwear pooled) obtained with mechanical test method on dry ice for both labs (IRSST and KITE)

Mode	N	Subsets for **IRSST** – COF **All Footwear**		
		1	2	3
Forepart	60	0.158		
Heel	60		0.182	
Flat	60			0.204

Mode	N	Subsets for **KITE** – COF **All Footwear**	
		1	2
Forepart	30	0.110	
Heel	30	0.111	
Flat	30		0.151

The effect of the mode did not seem to be the same from one boot to another, meaning that the ranking of the boots could be different depending on the mode of slipping (significant effect of interaction between *footwear* and *mode*, Table 4). The highest difference was with boot F10, assessed on the basis of the heel slip mode at the IRSST lab (see Table 27 in APPENDIX C). In this situation, the boot ranked 6th (subset 4/6), rather than 9th in the last (or second-to-last) subset, as was the case for the other modes. With this one exception, boots F1, F2, F3, F5 and F6, circled in red in Figure 12, generally had lower COFs than boots F4, F7, F8, F9 and F10, circled in blue.

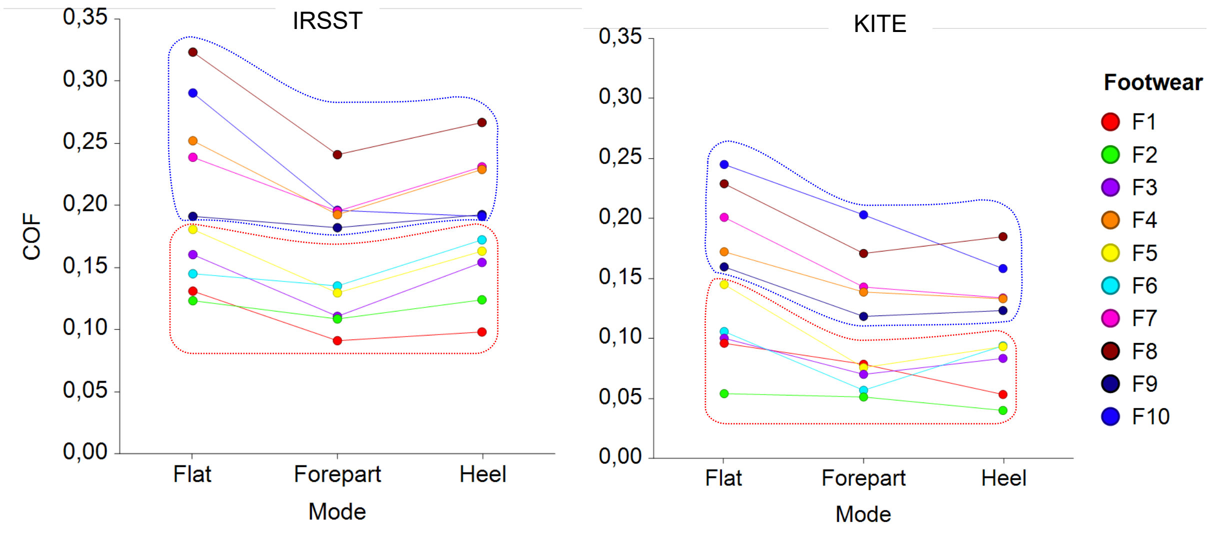
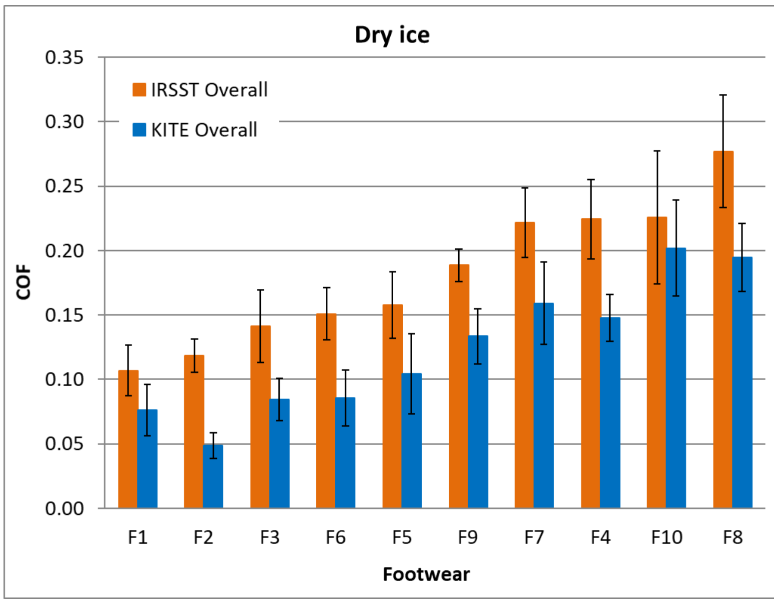
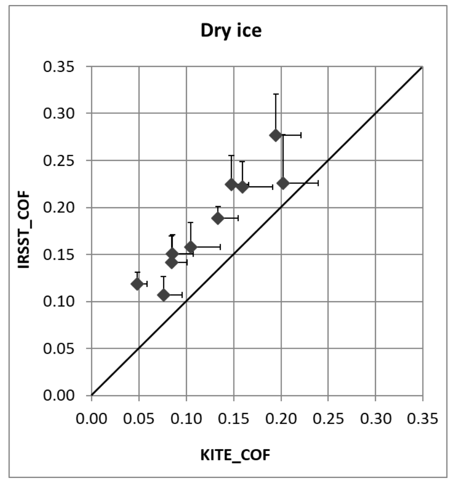


Figure 12. Footwear overall COF for each mode (flat, forepart, heel) obtained with the mechanical test method on dry ice for both labs (IRSST and KITE).

The chart comparing overall COFs ($\overline{COF_f}$) obtained at the two labs (Figure 13) shows that the values recorded at IRSST were systematically higher than those obtained at KITE. The B&A analysis (Figure 13) showed that the bias between the two labs was 0.058, which is fairly high. However, the analysis also showed that the level of agreement was fairly good (0.036). In addition, there was a strong correlation between the two labs (R = 0.94).





Bias	LoA	R
0.058	0.036	0.94

Figure 13. Comparison of IRSST vs. KITE overall COF results for 10 types of footwear, evaluated using the mechanical test method in dry ice conditions. In the graph on the left, footwear models are presented in ascending order of the results obtained at IRSST. In the graph on the right, the black solid line represents what would be a perfect agreement and correlation between IRSST and KITE results. The table summarizes the results of the B&A analysis, i.e., bias, limits of agreement (LoA) and correlation coefficient (R) between lab measurements. Detailed results are in APPENDIX C, numerical results of the left graph (Table 22), B&A analysis (Figure 27, Figure 28, and Table 26).

4.2.3.2 Wet ice

The ANOVAs performed for each laboratory for wet ice (Table 7) showed that the main factors (*footwear* and *mode*) and the interaction between these two factors explain most of the differences between the COFs measured at the two labs. A large effect size for the factor *footwear* (η_p^2 > 0.9) was observed. For the *mode* factor (and its interaction with the *footwear* factor), the effect size was less for the measurements taken at IRSST ($\eta_p^2 \approx 0.371$) than for those taken at KITE ($\eta_p^2 \approx 0.849$), as was the case for dry ice.

In Table 8, significant differences can be seen between the boots assessed by Tukey multiple comparisons. The $\overline{COF_f}$ values recorded at IRSST were generally of the same order of magnitude as those obtained at KITE for all boots. The subsets for the performance ranking were slightly different between the two labs. At the IRSST lab, the boots shown in red (F1, F2, F3, F5 and F6) are classified into 4 subsets, generally by groups of 3 boots per subset, thus revealing a few significant differences between certain models (between F5 and F3, between F5 and F6, and between F2 and F6). At the KITE lab, these same boots were classified in the same subset, thus indicating no significant difference between the 5 boots. Despite a few differences observed between the boots in the ranking (significance patterns), both the IRSST and KITE labs were able to determine the superior performance of boots F9 and F10, whose COF values were noticeably higher than those of the other boots and exceeded the threshold of 0.12.

Table 7. ANOVAs for mechanical test results on wet ice for both labs (IRSST and KITE), including effect of footwear and mode (significant with $p \leq 0.05$)

Lab	Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
IRSST	Footwear	.573	9	.064	413.751	<.001	.962
	Mode	.013	2	.007	43.853	<.001	.371
	Operator	.001	1	.001	3.517	.063	.023
	Footwear * Mode	.013	18	.001	4.633	<.001	.359
	Error	.023	149	1.537E-4			
	Total (corrected)	.622	179				
KITE	Footwear	.220	9	.024	250.605	<.001	.974
	Mode	.033	2	.016	169.324	<.001	.849
	Footwear * Mode	.014	18	.001	7.850	<.001	.702
	Error	.006	60	9.738E-5			
	Total (corrected)	.272	89				

Table 8. Footwear ranking and significant subsets according to overall COFs, obtained with mechanical test method on wet ice for both labs (IRSST and KITE)

Foot wear	N	Subsets for **IRSST** – Overall COF						
		1	2	3	4	5	6	7
F5	18	0.036						
F2	18	0.043	0.043					
F1	18	0.046	0.046	0.046				
F3	18		0.051	0.051	0.051			
F6	18			0.057	0.057			
F7	18			0.059	0.059			
F4	18				0.062			
F8	18					0.089		
F9	18						0.184	
F10	18							0.200

Foot wear	N	Subsets for **KITE** – Overall COF				
		1	2	3	4	5
F5	9	0.028				
F6	9	0.028				
F2	9	0.030				
F1	9	0.039				
F3	9	0.040				
F7	9		0.058			
F4	9			0.077		
F8	9			0.087		
F9	9				0.148	
F10	9					0.226

The significant differences between the modes of slipping, for all boots, are presented in Table 9. At both labs, the three modes of slipping produced significantly different COFs. The lowest COF values were generally measured in the heel slipping mode, and the highest in flat mode, with mean differences of approximately 0.02 for IRSST and 0.04 for KITE.

Table 9. Significant differences between modes according to COF (all footwear pooled), obtained with mechanical test method on wet ice for both labs (IRSST and KITE)

Mode	N	Subsets for **IRSST** – COF **All Footwear**		
		1	2	3
Heel	60	0.073		
Forepart	60		0.081	
Flat	60			0.094

Mode	N	Subsets for **KITE** – COF **All Footwear**		
		1	2	
Heel	30	0.059		
Forepart	30		0.070	
Flat	30			0.100

Given that a test method should be able to separate good-performance boots from ones that aren't as good, it can be seen that the mechanical method was able to do so on wet ice, regardless of the mode of slipping. Boots F9 and F10 recorded COFs that were higher than the other boots for all modes of slipping (Figure 14, and Table 28 in APPENDIX C).

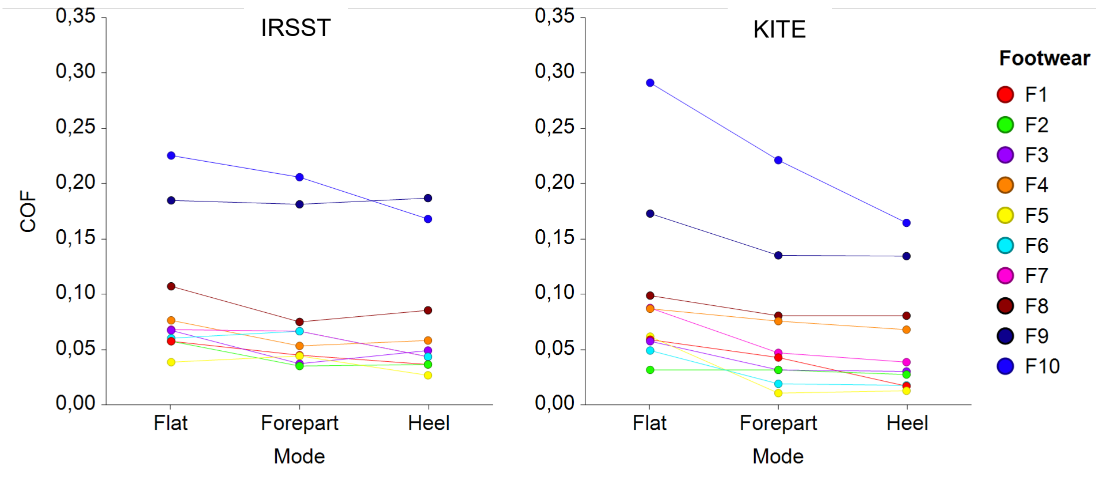
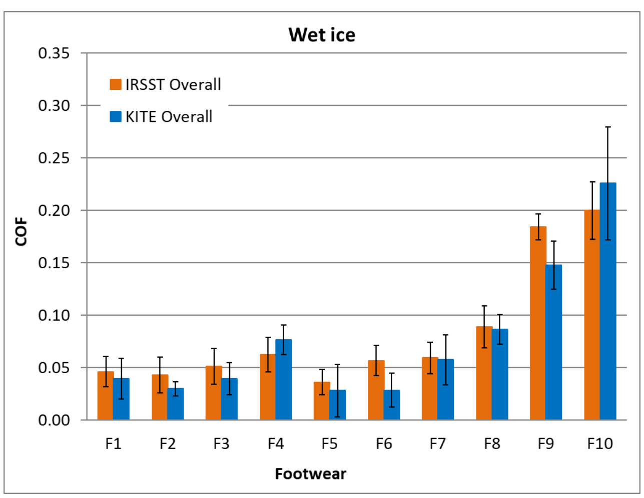
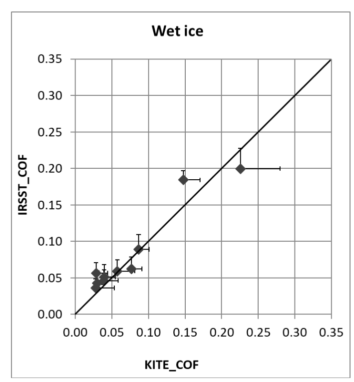


Figure 14. Footwear overall COF for each mode (flat, forepart, heel), obtained with the mechanical test method on wet ice for both labs (IRSST and KITE).

The chart comparing the overall COFs ($\overline{COF}_f$) obtained at the two labs (Figure 15, left graph) shows that the values were similar at IRSST and at KITE. The B&A analysis (Figure 15, right graph) indicates that the bias between the two labs was low, at 0.007. It also shows that the level of agreement was fairly good (0.036) and that there was a strong correlation between the two labs (R = 0.96).





Bias	LoA	R
0.007	0.036	0.96

Figure 15. Comparison of IRSST vs. KITE overall COF results for 10 types of footwear evaluated using the mechanical test method in wet ice conditions. In the graph on the left, footwear models are presented in ascending order of the results obtained at IRSST. In the graph on the right, the black solid line represents what would be a perfect agreement and correlation between IRSST and KITE results. The table summarizes the results of the B&A analysis, i.e., bias, limits of agreement (LoA) and correlation coefficient (R) between lab measurements. Detailed results are in APPENDIX C, numerical results of the left graph (Table 23), B&A analysis (Figure 27, Figure 28, and Table 26).

4.3 Discussion

One of the objectives of phase 1 was to develop an alternative mechanical test method based on methods ASTM F2913-11 and SATRA TM144:2011, on dry and wet ice surfaces similar to those of the WinterLab. To develop this mechanical method, certain choices had to be made in the test protocol (ice making, experimental unit). The other objective of phase 1 was to assess the repeatability and reproducibility of the method by comparing the results obtained for 10 winter work boot models in two labs.

Intra-laboratory variability

The **intra-operator repeatability** of the mechanical method used in this study was generally acceptable for all operators (IRSST and KITE), with pooled SDs of 0.014 for dry ice and of 0.009 for wet ice (Table 18 and Table 19, APPENDIX B). The **inter-operator reproducibility** (IRSST) was satisfactory on the whole, with pooled SDs of 0.018 for dry ice and of 0.013 for wet ice (Table 20, APPENDIX B). This result was also confirmed by the ANCOVA, which showed that the trained operator did not have any significant effect on the measurements (Table 3). These measures can be put into perspective by comparing them to a shoe-on-quarry tile inter-laboratory study of standardized test method ASTM F2913-19 (ASTM International, 2019b), involving 10 labs and focusing on a single model of shoe, where for the heel slip mode, repeatability SD of 0.020 on a wet quarry tile and of 0.044 on a dry quarry tile, were obtained. Reproducibility SD of 0.046 on a wet quarry tile and of 0.086 on a dry quarry tile were also reported in this former study. It should be noted that the mean COF of the boot tested (0.513 on a wet quarry tile and 0.564 on a dry quarry tile) was higher than those recorded in this study on ice (COF < 0.277), given that the quarry tiles are less slippery than ice surfaces.

In addition, the current study found that the mechanical method yielded CVs generally less than 10%. This was the case for two-thirds of the tests on dry ice, but for less than half of the tests on wet ice. It is likely that the lower intra-operator repeatability and inter-operator reproducibility on wet ice were due in part to the low resolution of the data acquisition system of the STM 603 device, which is 0.01. A resolution of 0.001 would be preferable in cases where low COFs are measured (as was the case for boots F1 to F8, for which the COF < 0.11).

To improve the mechanical method's precision, it would be useful to minimize the sources of variability. Though the laboratory environment did not have a significant effect on the COFs measured (Table 3), despite the relatively large temperature and relative humidity ranges in the testing, it would be preferred if the testing laboratories had temperature and relative humidity control systems. That would further improve the repeatability of the ice making conditions and allow testing to occur year round under controlled temperature and humidity conditions.

Reconsidering the choice of experimental unit (mean of last 5 consecutive runs) may be warranted. Although an *a posteriori* assessment showed that the variation in the last 5 consecutive runs was 10% or less in most cases out of the 540 test results (Table 17), some cases did reveal variations exceeding 10% over successive runs, especially on wet ice. More in-depth study of these variations in successive runs would be valuable, especially as the COF values are heavily influenced by the ice refrigeration cycle, which can show differing rates of change over the cycle (Figure 7). It should be noted that this experimental unit is not suited to soles with crampons or grit, like boot F9 with Green Diamond technology. For that boot, the COF of the first test run would probably have been more representative of the sole slip resistance compared to the mean of the five measurements taken when the grit had gone through the same grooves several times.

Inter-laboratory variability

Regarding inter-laboratory reproducibility, the analyses showed that the measurements with **wet ice** were equivalent in both labs with respect to both similar COF values obtained (the bias between the two labs was just 0.007) and similar ranking of the boots (in which boots F9 and F10 stood out as distinctly better performers than the other models). In addition, the ANOVA showed that, for this type of ice, the laboratory effect was relatively weak (η_p^2 = 0.069, Table 25). The agreement between the two labs was relatively good (LoA = 0.036) and the linear correlation was high (R = 0.96).

For **dry ice**, the COF values recorded at IRSST were systematically higher than those obtained at KITE, with a bias of 0.058 for the overall COFs. Still, the agreement between the two labs was relatively good (LoA = 0.036) and the linear correlation was high (R = 0.94). Furthermore, the ranking of the boots was more or less equivalent, with both labs identifying boots F4, F7, F8, F9 and F10 as significantly better performers than boots F1, F2, F3, F5 and F6. The only less consistent result was for boot F10, assessed on the basis of the heel slip mode at the IRSST lab, which ranked 6th rather than 9th or 10th, as was the case for the flat and forepart modes at IRSST and for all modes at KITE (Table 27 in APPENDIX C).

There are a number of sources of variability that may explain the differences noted between the two labs. First, the refrigeration cycles showed different patterns, especially for dry ice (Figure 5). For this type of ice, the test windows had to be defined at a time when the ice was warming up at IRSST (from -2°C to -1°C) and cooling down at KITE (from -1°C to -2°C) (Table 2). The differences in patterns observed in the refrigeration cycles could be due to possibly different operation of the ice trays (among other things, the warmer and colder areas are not located at the same places on the two trays). They may also be due to the laboratory environment, where the temperature, the relative humidity and the speed of circulation of the ambient air could noticeably affect thermal exchange between the air and the ice surface and have an impact on the ice refrigeration cycle.

Another source of variability can be found in the procedure for removing frost from the ice surface with a wet cloth. This procedure was performed at the start of each series of tests. While meetings were held with operators from the two labs to standardize the procedure (same cloth, approximately same amount of water), a better match between the procedures would be desirable, such as the length of time that elapses between when the wet cloth was used and the tests were run, which was not exactly the same between IRSST (~10-30 seconds before the tests) and KITE (1-2 minutes before the tests).

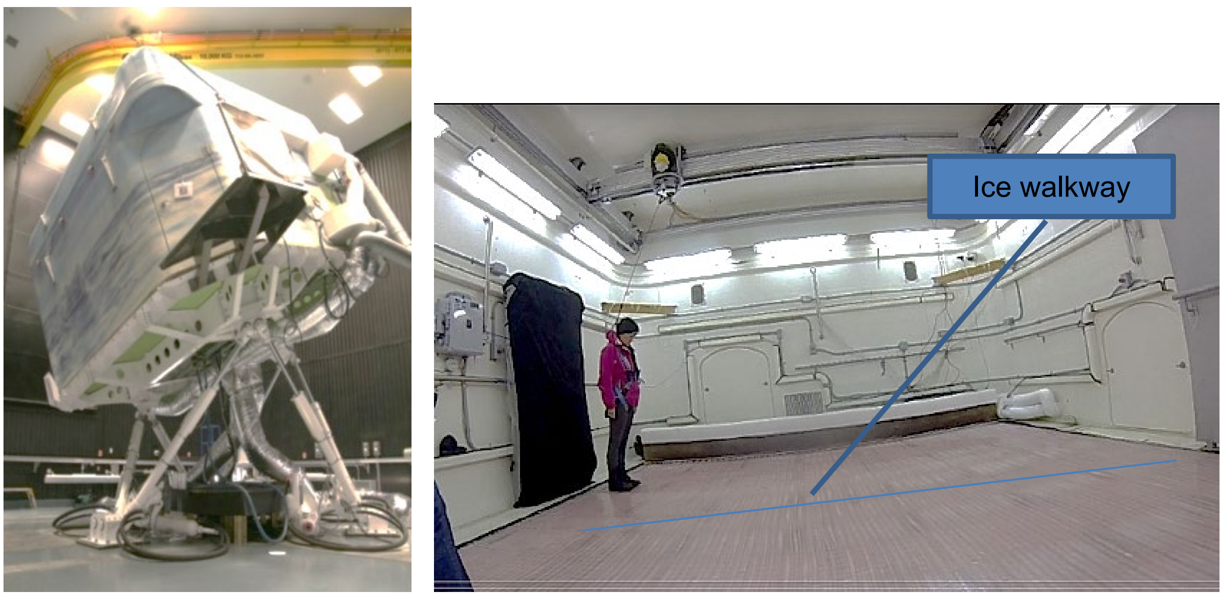
Generally speaking, greater standardization of the entire test procedure at the two labs would be useful, as would improving control over the ambient conditions and the ice surface.

5. PHASE 2 – COMPARISON WITH HUMAN-CENTRED MAA METHOD

5.1 Method

5.1.1 Apparatus

The human-centred MAA method, developed by KITE, was used to evaluate footwear slip resistance on ice surfaces. The WinterLab (Figure 16a) contains a 2.5 cm thick, 4.5 m × 4.5 m floor (Figure 16b) that can accommodate varying temperatures for each surface condition. Different angles can be created in the WinterLab using a hydraulic-powered motion base to tilt the entire walkway surface (Figure 16a). Participants walked on a 4.0 m × 0.9 m pathway in the centre of the WinterLab while wearing a full-body overhead safety harness attached to a fall-arrest device.



(a) (b)

Figure 16. (a) WinterLab mounted on a hydraulic-powered motion base to create slopes; (b) Inside the WinterLab (ice walkway is marked by lines embedded in the ice).

5.1.2 Ice conditions

The tests were conducted on two different ice surfaces.

Dry ice: The entire WinterLab floor surface was flooded with water and cooled to a temperature of -5.5 ± 1°C (Bagheri, Patel, Li, Morrone, et al., 2019; Bagheri, Patel, Li, Rizzi, et al., 2019). Ice temperature was controlled using tubes circulating glycol coolant along the floor surface, and an ambient air temperature of 2.5 ± 2°C was maintained throughout the experiment. In combination, these factors allowed a smooth, dull, ice walkway surface to be formed with minimal melting at the interface (i.e., with no water visible on the ice surface). The relative humidity in the WinterLab for dry ice conditions was about 45%.

Wet ice: Starting with a dry ice surface, wet ice conditions were created by holding the ambient air temperature of the room at 8.0 ± 2°C and the ice surface temperature at -1.5 ± 1°C (Bagheri, Patel, Li, Morrone, et al., 2019; Bagheri, Patel, Li, Rizzi, et al., 2019). The warmer ambient temperatures in combination with the near freezing ice temperature helped to maintain a thin layer of water over the ice surface. The relative humidity in the WinterLab for wet ice conditions was about 36%.

5.1.3 Procedures and test participants

Different surface angles were created in the WinterLab using KITE's single axis motion base to tilt the entire walkway surface. Starting with a level surface at 0°, each participant was required to walk at a self-selected pace up and down the incline surface. The slope angle was progressively increased by 1° increment until the failure angle was reached. An angle was considered to be the failure angle if the participant could not initiate gait or if both feet slipped simultaneously while traversing the floor. The participant had to slip on two separate trials before an angle was considered to be the failure angle. Likewise, the participant had to walk successfully on two separate trials at the immediately preceding angle. This angle is deemed the "Maximum Achievable Angle" (MAA). The higher the MAA, the better the slip resistance.

Four female participants were recruited to walk with women's footwear (F2, F3, F6 and F7, see Table 1), and four male participants were recruited to walk while wearing men's footwear (F1, F4, F5, F8, F9 and F10, see Table 1) in random order. The criteria used in sample size estimation for the MAA method was that a 95% confidence interval (CI) for the mean maximum achievable angle for each type of footwear should be no more than two degrees in total width. In other words, the sample size and the number of repetitions that will yield a 95% CI of x ± 1.0° had to be determined. An internal study had been conducted at KITE in 2014 with 15 male and 15 female participants to establish the required sample size. The results revealed that a sample size of 4 with a repetition of 1 yielded a 95% CI of x ± 0.95°, and a sample size of 8 with a repetition of 3 yielded a 95% CI of x ± 0.48°[4]. Therefore, for most of the MAA test, a sample size of 4 is considered as a good compromise between cost efficiency and accuracy.

The participants were between the ages of 19 to 28 years and had no known musculoskeletal dysfunctions or mobility limitations. The average age was 23.3 ± 1.1 years, average height 165.7 ± 4.9 cm and average weight 64.4 ± 8.8 kg for the 4 female participants. The average age was 23.0 ± 3.7 years, average height 177.3 ± 7.8 cm, and average weight 70.8 ± 11.1 kg for the 4 male participants. Participants were asked to walk naturally throughout the walking tests. Although not specifically trained, participants performed 1-2 familiarization trials before the test so that they could be familiar with the WinterLab experiments. The MAA test protocol was approved by the KITE ethics committee to ensure the safety of the experiments and data privacy. Participants signed a consent form before participating in the study.

Participants were asked to select the boot size that fit them the best. Three pairs of each type of footwear (US sizes 9, 10 and 11 for men's models and US sizes 7, 8 and 9 for women's models)

[4] A sample size analysis completed by KITE statistical consultant Dr. Ellen Maki (Analytica Statistical Consulting Inc., Toronto) based on previous MAA testing data indicated that footwear tested with 4 participants provides sufficient power to find statistically significant differences with 95% confidence for pair-wise comparisons of any footwear that had MAA differing by 2° or more.

were used so that at least one participant used the same pair of boots as another participant wearing the same size. Participants were informed of this situation.

5.1.4 Testing conditions

Using the MAA test method, the 10 types of winter footwear were tested by 4 participants on two ice surfaces (dry and wet) in both uphill and downhill directions (Figure 17), giving a total of 160 MAAs.

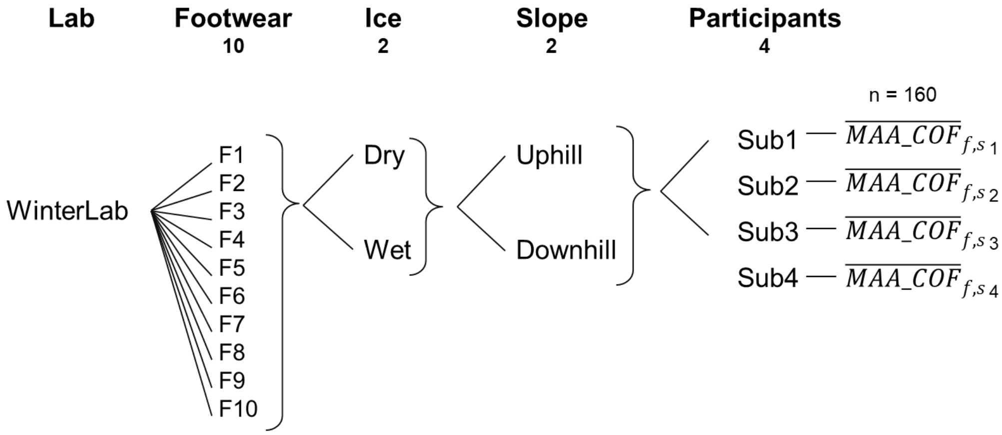


Figure 17. MAA testing conditions and frequency.

The boots were all brand new. The preparation of the footwear sole was the same as for the mechanical testing, i.e., rub with silicon carbide paper and wash with detergent solution. The footwear was conditioned for 3 hours inside the WinterLab before testing.

5.1.5 Analyses

For each ice condition (dry and wet), each footwear model (F1 to F10), and each slope direction (uphill and downhill), four MAAs were recorded (for the four participants). Each MAA was converted to COF values (MAA_COF) by taking the tangent of the angle:

$$MAA\_COF = \tan(MAA) \quad \textbf{(5)}$$

The average MAA_COF over participants (assuming that the MAAs have Gaussian distribution), standard deviation and coefficient of variation were then calculated for each ice condition, using the following equations:

$$\overline{MAA\_COF_c} = \frac{1}{n}\sum_{i=1}^{n} MAA\_COF_{c,i} \quad \textbf{(6)}$$

$$\boldsymbol{SD_c} = \sqrt{\frac{\sum_{i=1}^{n}\left(\boldsymbol{MAA\_COF_{c,i}} - \overline{\boldsymbol{MAA\_COF_c}}\right)^2}{(\boldsymbol{n}-\boldsymbol{1})}} \quad \textbf{(7)}$$

$$\boldsymbol{CV}\ (\%) = \frac{\boldsymbol{SD_c}}{\boldsymbol{\overline{MAA\_COF}_c}} * \mathbf{100} \tag{8}$$

where c represents the testing conditions considered, i the number of the test result and n the total number of test results, as in the following:

- $\boldsymbol{c} = \boldsymbol{f,s}$ for each footwear model ($\boldsymbol{f}$ = ***F1, F2, … F10***), each slope direction ($\boldsymbol{s}$ = ***up, down***), with $\boldsymbol{n}$ = 4 (giving 20 $\overline{MAA\_COF}_{f,s}$ for each ice condition)

- $\boldsymbol{c} = \boldsymbol{f}$ for each footwear model ($\boldsymbol{f}$ = ***F1, F2, … F10***), both slope directions combined, with $\boldsymbol{n}$ = 8 (giving 10 $\overline{MAA\_COF}_f$ for each ice condition)

To facilitate the interpretation and analysis of boot performance, similar to phase 1, the obtained $\overline{MAA\_COF}_f$ were compared with the threshold of 0.12, defined by one snowflake. In addition, the boot performances were classified into the same two subsets of 5 boots each, depending on whether they obtained the highest (in blue) or lowest (in red) $\overline{COF}_f$, in accordance with the mechanical method on dry ice.

A three-way analysis of variance (ANOVA) with the factors of ice surface (dry, wet), footwear (F1 to F10) and slope direction (uphill, downhill) was used to determine their effects on MAA_COF. Then, Tukey multiple comparison tests were done to rank the performances of the boots by significantly different subsets, as well as for the directions. Statistical analyses (ANOVA and Tukey) were carried out using SPSS statistical software version 23 (IBM Corporation, 2015), at a significance level of 0.05. The assumptions underlying the use of the models were verified by examining the model's standardized residuals.

The **comparison between the human-centred MAA and the mechanical test methods** was made using Bland-Altman (B&A) analyses, as well as by comparing the footwear performance rankings (Bland & Altman, 2010). The mechanical test results from the two labs, IRSST_COF and KITE_COF, taken separately, were compared with the MAA_COF values. As in phase 1 analyses, the B&A analyses served to assess the bias and the limits of agreement (LoA) where 95% of the differences between the measurement series (IRSST_COF vs. MAA_COF, KITE_COF vs. MAA_COF) are found, as well as the linear correlation coefficient (R) between the measurement series. Series of measurements are considered to be in agreement if the bias is almost nil (close to zero), the limits of agreement are very low, and the coefficient of correlation is close to 1.0.

These B&A analyses were performed with NCSS 11 Statistical Analysis Software (Hintze, 2004; NCSS Statistical Software, 2016), using $\overline{MAA\_COF}_f$ and $\overline{COF}_f$ (Design 1 comparison in NCSS 11). Analyses were also done with $\overline{MAA\_COF}_{f,s}$ and $\overline{COF}_{f,m}$ taking into account multiple repetitions for slope directions and modes of slipping (Design 2 comparison in NCSS 11).

5.2 Results

The ANOVA (Table 10) shows that the main factors of *footwear*, *ice* and *slope*, as well as the interaction *footwear*ice* had a significant effect on the measurement of MAA_COF ($p < 0.001$). *Footwear* is the factor that has the greatest effect (effect size $\eta_p^2 \approx 0.9$), which is desirable, given that the MAA test method serves specifically to measure, compare and rank the performance of different boots. Effect sizes were less for the factors *ice* (η_p^2 ≈ 0.1) and *slope* (η_p^2 ≈ 0.2). The existence of a significant interaction between the factors *footwear* and *ice*, with a relatively strong effect size (η_p^2 = 0.610), indicates that the effect of the boots was not the same on the two ice surface conditions.

Table 10. ANOVA for MAA_COF results, including effect of footwear, ice surface and slope direction (significant with $p \leq 0.05$)

Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
Footwear	.349	9	.039	94.876	<.001	.877
Ice	.006	1	.006	15.412	<.001	.114
Direction	.009	1	.009	21.384	<.001	.151
Footwear * Ice	.077	9	.009	20.840	<.001	.610
Footwear * Direction	.004	9	4.645E-4	1.137	.342	.079
Ice * Direction	.001	1	.001	2.975	.087	.024
Footwear * Ice * Direction	.001	9	1.650E-4	.404	.931	.029
Error	.049	120	4.083E-4			
Total (corrected)	.496	159				

The significant differences between slope directions, presented as subsets in Table 11, were assessed by means of Tukey multiple comparisons. Globally, walking uphill achieved higher MAA_COF, especially on dry ice. For dry ice, the MAA_COF values were 0.02 (1.1°) higher for uphill than for downhill. For wet ice, the difference (not significant) between downhill and uphill was less than 0.01 (0.5°).

Table 11. Significant differences between downhill and uphill slope directions for MAA_COF of all footwear and ice conditions pooled (left), and for dry (middle) and wet (right) ice conditions (all footwear pooled)

Direction	N	Subsets	
		1	2
Downhill	80	0.016	
Uphill	80		0.106

Direction	N	Subsets for **dry ice**	
		1	2
Downhill	40	0.095	
Uphill	40		0.115

Direction	N	Subset for **wet ice**
		1
Downhill	40	0.088
Uphill	40	0.097

The significant differences between the boots are shown in Table 12. The MAA method classified the boots into three subsets.

The ranking of the boots by the MAA method for **wet ice** (Table 12) is similar to that done by the mechanical method, with a few exceptions. Using the MAA method, it was found that boot F1 had a significantly lower COF than most of the other boots, including F3, F5 and F6, whereas the mechanical method did not reveal any significant differences between these four models of boot (Table 8). Also, boots F7, F8, and F4 (in blue) had results similar to those of boots F2 and F5 (in red) with the MAA method, but not with the mechanical method (where F7, F4, F8 > F2, F5). Aside from these ranking differences, which are within a relatively narrow COF interval (0.026 to 0.089), the MAA and mechanical methods both agreed on the higher COF values for boots F9 and F10 (COFs > 0.12). Therefore, on wet ice, both mechanical and MAA methods gave similar assessments to the boots that stood out favourably from the others.

Table 12. Footwear ranking and significant subsets according to MAA COF on dry (left) and wet (right) ice conditions

Foot wear	N	Subsets for **MAA** COF **dry ice**		
		1	2	3
F4	8	0.072		
F2	8	0.081	0.081	
F5	8	0.088	0.088	
F7	8	0.088	0.088	
F1	8	0.090	0.090	
F8	8	0.094	0.094	
F6	8		0.112	
F3	8		0.112	
F10	8			0.158
F9	8			0.158

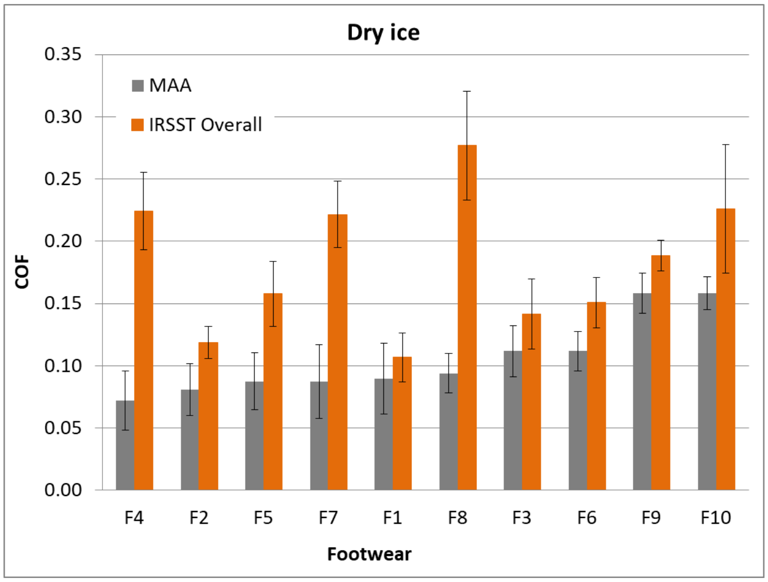
Foot wear	N	Subsets for **MAA** COF **wet ice**		
		1	2	3
F1	8	0.026		
F2	8	0.046	0.046	
F3	8		0.066	
F4	8		0.066	
F5	8		0.070	
F7	8		0.072	
F6	8		0.072	
F8	8		0.074	
F9	8			0.174
F10	8			0.261

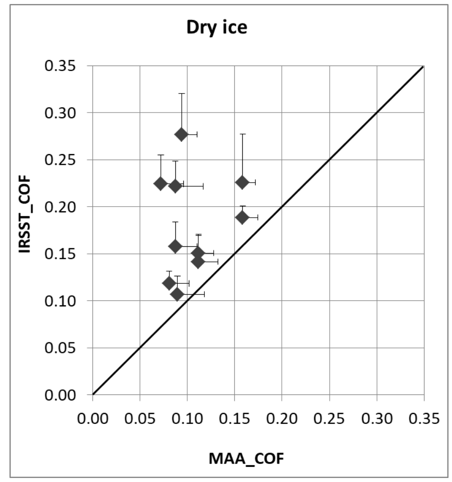
For **dry ice**, the boot ranking showed notable differences between the two methods. First of all, the MAA method found virtually no significant differences between boots F1 to F8. For instance, the ranking results show that the boot F4 had the lowest COF (0.072), while boot F8 has a slightly higher COF (0.094), with no significant differences between the two boots. The mechanical method not only assessed significant differences between these same boots, but also rated boots F4 and F8 among the best performers with COFs greater than 0.12 (Table 5). The MAA method rated F9 and F10 as having COFs that were distinctly higher than those of the other boots (Table 12). For these same boots, F9 and F10, the mechanical method also provided higher COFs.

The chart comparing the COFs obtained from the MAA method ($\overline{MAA\_COF_f}$) with those obtained from the mechanical method ($\overline{COF_f}$) on **dry ice** (Figure 18) also shows substantial differences. With the MAA method, only boots F9 and F10 had high COFs, exceeding the threshold of 0.12. With the mechanical method, on the other hand, several other boots also had high COFs, not only exceeding the threshold of 0.12, but also revealing some performance results inconsistent with those assessed with the MAA method. For example, boots F4, F7 and F8 (both for IRSST and KITE) showed good resistance to slipping, but they did not stand out when tested with the MAA method.

The B&A analysis (Figure 18) confirmed that the two methods were not in agreement for **dry ice**. The limits of agreement were fairly high (LoA = 0.135 for IRSST and 0.117 for KITE) and the linear correlation between the two methods was low (R = 0.14 for IRSST and 0.34 for KITE). The bias between the two methods was higher for the IRSST lab (0.076) than for the KITE lab (0.018).

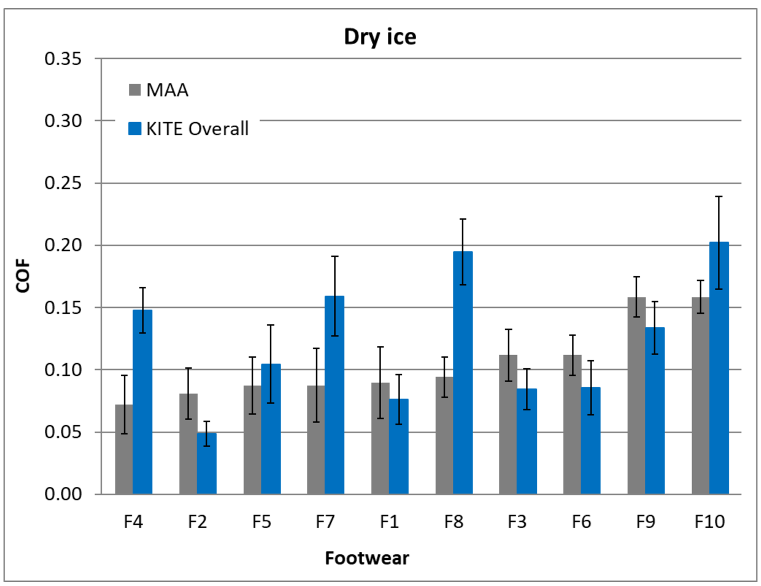
For **wet ice**, however, the mechanical method seems overall to be in agreement with the MAA method (Figure 19). The two methods show better performances for boots F9 and F10, which stand out from the performances of the other boots. In addition, the B&A analysis (Figure 19) confirms that the two methods are in agreement. The limits of agreement were fairly low between the two methods (0.070 for IRSST and 0.074 for KITE), the biases were low (-0.010 for IRSST and -0.017 for KITE) and the linear correlation was good (0.95 for IRSST and KITE). However, there was a cluster of data points with low COFs, and two points with high COF (Figure 19). Therefore, the correlation between the two methods on wet ice should be interpreted carefully because there was not a substantial COF gradation in the tested boots.

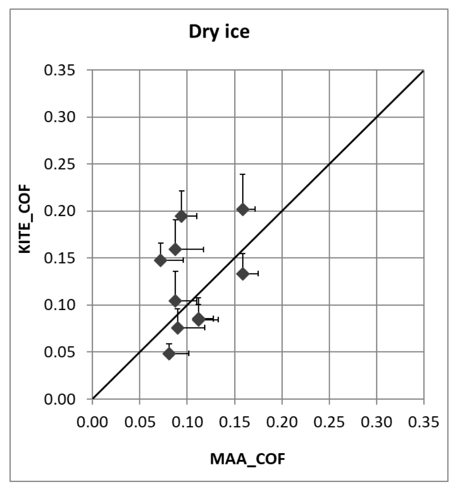




Bias	LoA	R
0.076	0.115	0.14

(a)

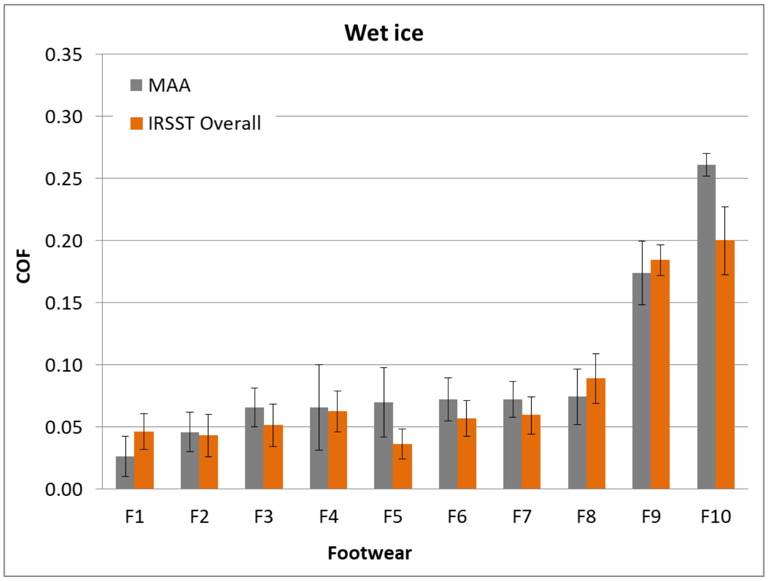


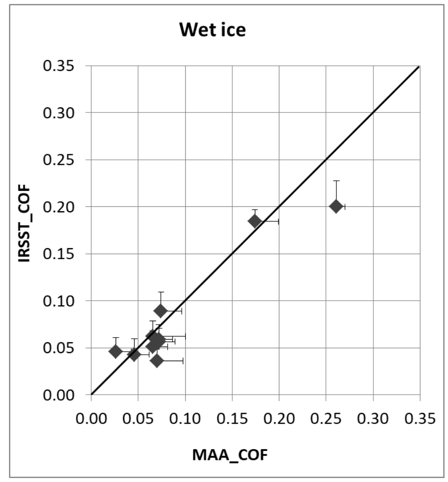


Bias	LoA	R
0.018	0.099	0.34

(b)

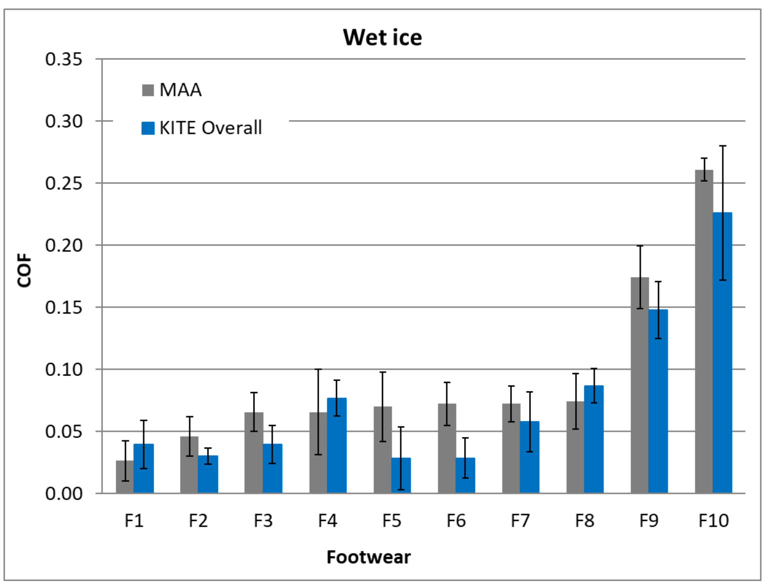
Figure 18. MAA COF results for 10 types of footwear on dry ice conditions and comparison with overall COF at IRSST (a) and at KITE (b) labs using the mechanical test method. In the graphs on the left, footwear models are presented in ascending order of MAA results. In the graphs on the right, the black solid line represents what would be a perfect agreement and correlation between COF and MAA_COF. The tables summarize the results of B&A analyses, i.e., the bias, limits of agreement (LoA) and correlation coefficient (R) between COF and MAA COF measurements. Detailed results are in APPENDIX D, with numerical results of the left graph in Table 22, and B&A analysis in Figure 29 and in Table 29.

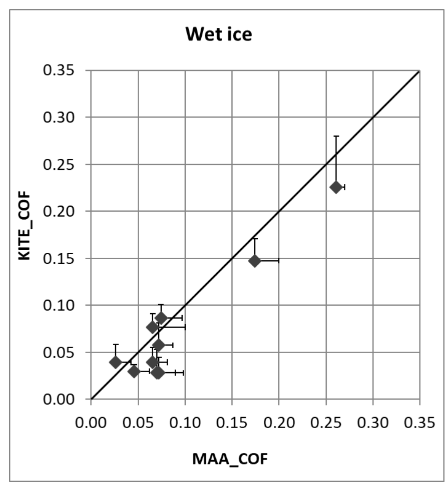




Bias	LoA	R
-0.010	0.047	0.94

(a)





Bias	LoA	R
-0.017	0.043	0.95

(b)

Figure 19. MAA COF results for 10 types of footwear on wet ice conditions and comparison with overall COF at IRSST (a) and at KITE (b) labs using the mechanical test method. In the graphs on the left, footwear models are presented in ascending order of MAA results. In the graphs on the right, the black solid line represents what would be a perfect agreement and correlation between COF and MAA_COF. The tables summarize the results of B&A analysis, i.e., the bias, limits of agreement (LoA) and correlation coefficient (R) between COF and MAA COF measurements. Detailed results are in APPENDIX D, numerical results of the left graph in Table 23, B&A analysis in Figure 30 and in Table 29.

5.3 Discussion

The goal of phase 2 was to compare the mechanical test method with the MAA human-centred method. The comparison was made on the basis of the agreement between the methods (bias, limits of agreement and correlation) and footwear ranking.

The COFs and boot ranking obtained on **wet ice** with the mechanical method were close to those obtained with the MAA method. The B&A analysis showed a good agreement between the two methods at both IRSST and KITE labs. Both methods also produced similar boot rankings, with boots F9 and F10 obtaining distinctly higher COFs than the other boots with COF higher than the threshold of 0.12. However, the fairly good correlation between the two methods on wet ice should be carefully interpreted. More footwear with moderate slip resistance on wet ice need to be tested to conclude on the correlation between the two methods.

On **dry ice**, however, the mechanical method did not show a good agreement with the MAA method and its boot ranking was quite different. The MAA method found virtually no significant difference between boots F1 to F8 and ranked them as the lowest performing, under the threshold of 0.12. In contrast, the mechanical method (at both IRSST and KITE) not only found significant differences between many of these boots, but also rated boots F4, F7 and F8 as being among the best performers, exceeding the threshold of 0.12 (Table 5, Figure 18). It may therefore be possible that the mechanical method overestimates the performance of some boots in certain cases. It is worth noting, nevertheless, that the mechanical method, like the MAA method, found that the boots F9 and F10 offered good slip resistance.

Other studies, besides considering the ranking, have also used a method's ability to distinguish between different conditions as a method assessment criterion (Hsu et al., 2015; Powers, Blanchette, Brault, Flynn, & Siegmund, 2010). As for the mechanical method, it seems to better distinguish between the boots than the MAA method does (greater number of significantly different subsets). However, it is impossible to say whether this ability of the mechanical method to discriminate really reflects a performance difference that could be confirmed in the field. Studies have shown that boots that have performed better under the MAA method have indeed performed better in the field, under real winter conditions (Bagheri et al., 2020).

A number of things can explain the differences seen between the two methods. First of all, the differences may be explained by the fact that the test parameters used in the mechanical method do not reflect the human gait, as is the case of several mechanical methods (Chang, Grönqvist, Leclercq, Myung, et al., 2001; Chang, Grönqvist, Leclercq, Brungraber, et al., 2001; Grönqvist, Abeysekera, et al., 2001). To be comparable with the MAA method, the mechanical method used here would have to be improved by changing some of the test parameters, such as vertical force, slipping speed and heel contact angle, as other studies have done (Blanchette & Powers, 2015; Beschorner et al., 2019; Hunwin et al., 2010). These researchers have found values specific to these parameters in order to reduce the bias seen between the results of mechanical method ISO 13287 (2019), based on the STM 603 device, and the results obtained by an approach with human subjects.

In addition, although efforts were made to obtain ice surfaces with the ice tray that were as similar as possible to the ice surfaces of the WinterLab, the temperature at the ice surface, and especially the ambient air conditions were still not the same. Also, the conditioning of the boots prior to the tests was not done at the same temperature (cold temperature in the WinterLab, warm temperature for mechanical tests). However, the hardness of the soles, measured at different temperatures (APPENDIX A), was generally constant between 0°C and 24°C (Figure 26). It is therefore possible that this parameter did not have such a big impact on the difference in the results between the two test methods. The decision not to condition the boots at a cold temperature, for the mechanical method developed in this study, was made firstly because keeping the boot cold is quite a challenge in a laboratory where the temperature is 23°C and the temperature of the boot sole is declining at a rate of approximately 1°C per minute (APPENDIX A). Secondly, conditioning to cold does not allow doing as many repetitions in one day as conditioning at the laboratory ambient temperature. For the further development of the mechanical method, however, it could be useful to redo some tests by conditioning the boots at cold temperature.

It may also be worthwhile investigating the performance of the boots under real conditions, in the field, to get a better assessment of the value of the mechanical method in the laboratory and under a variety of real winter conditions.

6. PHASE 3 – LEVEL WALKING TEST METHOD FOR FOOTWEAR WITH CONFLICTING RANKINGS

A third pilot method was used as a “field test” to decide which of the two methods (mechanical or MAA) was the most reliable for evaluating footwear performance on ice, in cases where the results of phase 2 show that the mechanical and the human-centred test methods produced conflicting results. This method was developed and piloted by the KITE research team and consists in a human-centred level walking test method using a passive motion tracking system.

There was no formal sample size estimation done for this phase as this protocol was treated as a pilot study. To the best of our knowledge, this protocol was novel and consequently there was no previous data available to provide estimates for use in a sample size calculation. The number of participants used in this phase was set at five in order to exceed the number used in our MAA testing, which is based on four participants for each boot.

6.1 Method

6.1.1 Selecting footwear and ice surface

The test surface used for phase 3 was wet ice, as there is a greater risk of people slipping on wet ice than on dry ice. The two winter boots selected were F1 and F5, which had conflicting rating outcomes on wet ice conditions (summarized in Table 13). The results from the mechanical tests indicated that F1 and F5 were not statistically different in terms of slip resistance. MAA_COF, on the other hand, showed that the two boots had significantly different slip resistance qualities (angle of 1.7° versus 4.0°).

Table 13. Footwear information for phase 3: Overall COF results on wet ice

Footwear	IRSST				KITE				MAA	
	Overall	Heel	Forepart	Flat	Overall	Heel	Forepart	Flat	COF	angle
F1	**0.05**	0.04	0.04	0.06	**0.04**	0.02	0.04	0.06	**0.03**	1.7°
F5	**0.04**	0.03	0.04	0.04	**0.03**	0.01	0.01	0.06	**0.07**	4.0°
Comparison	F1 ≈ F5				F1 ≈ F5				F1 < F5	

6.1.2 Apparatus

A passive motion tracking system consisting of 1 Kestrel and 14 Raptor-E cameras, made by Motion Analysis (Santa Rosa, California), was installed in the WinterLab. The system recorded the position of markers attached to the footwear. Markers were placed at the toe, midfoot and heel areas of the footwear, as shown in Figure 20. The signals were collected at 150 Hz and passed through a Butterworth filter (4th order, zero-phase, 12 Hz cut-off frequency). Figure 21 shows one participant walking on wet ice wearing the two models of boot in the WinterLab.

Figure 20. Motion capture markers attached to footwear F1 (left) and F5 (right).

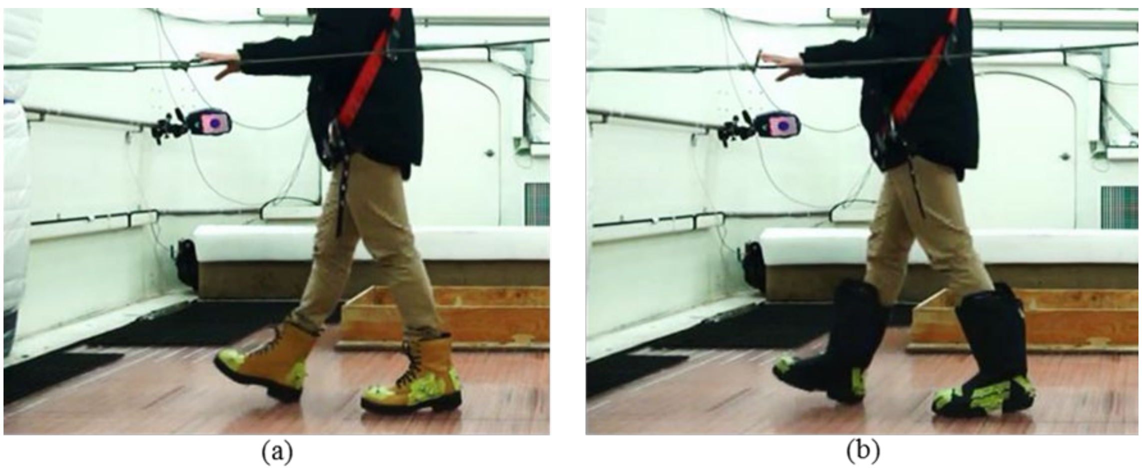


Figure 21. A participant walking with the (a) F1 and (b) F5 boots.

6.1.3 Test participants and procedures

Five healthy male participants were recruited to walk on a wet ice surface wearing the two selected types of occupational footwear. The participants were between the ages of 20 to 41 years and had no known musculoskeletal dysfunctions or mobility limitations. The average age was 28.0 ± 8.6 years, average height 174.2 ± 5.4 cm and average weight 65.2 ± 9.8 kg. Participants performed 1-2 familiarization trials before the test so that they could be familiar with the WinterLab experiments. They were instructed to walk back and forth along a linear path at a pace of 90 steps per minute using an auditory metronome, for five minutes. For safety, participants wore a fall arrest harness and were instructed to hover one of their hands above a rope and to only use the rope when they lost balance, as shown in Figure 21. Three participants began the test with boot F1, the two others with boot F5.

6.1.4 Analyses

MATLAB scripts written by KITE researchers were used to extract the steps stored in motion capture data by detecting the heel contact and toe-off from the velocity signal. Then, a previously developed machine learning algorithm (Cen, 2018) was applied to count the number of both forward heel and backward toe slips.

Slips at the heel were defined as the distance the foot travelled beyond 3 cm after heel contact, based on the study of Leamon and Li (1990). Therefore, 3 cm was adopted in this report as the threshold for counting heel slips. Since a fall is likely to happen on heel slips exceeding a distance of 10 cm, according to Strandberg and Lanshammar (1981), the number of hazardous heel slips was also counted. For toe slips, to the best of our knowledge, there was no previous study that considered slip distance for slips at the toe. Therefore, all toe slips after toe-off were counted.

Figure 22 shows samples of heel and toe slips captured by the markers. This figure also shows normal heel contact and toe-off for comparison. The algorithm was able to estimate slip distances with accuracies of 3 ± 13% and 3 ± 5% for heel and toe slips, respectively (Cen, 2018).

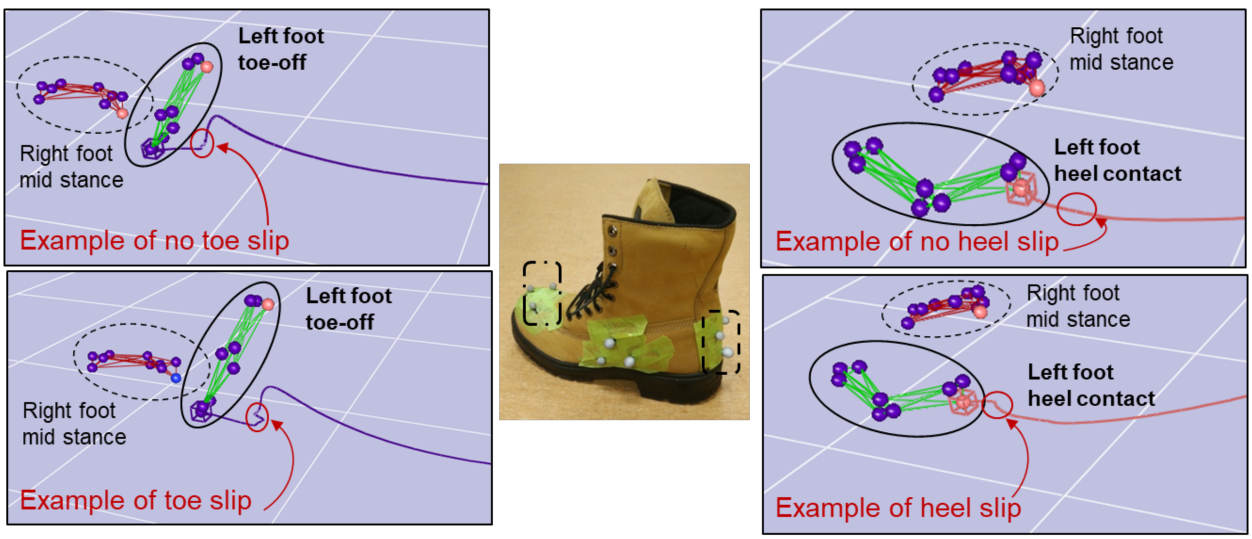


Figure 22. The upper images show normal toe-off and heel contact. The lower images show toe and heel slip samples.

The footwear performance was based on four parameters:

- Number of heel slips for which heel slip distance was > 3 cm after heel contact
- Number of toe slips for which toe slip distance was > 0 cm after toe-off
- Number of hazardous heel slips for which heel slip distance was ≥ 10 cm
- Participants' perception of the footwear's slipperiness, obtained from interviewing the participants after the trials with the question "which boot felt more slippery?"

The number of slips was calculated as a percentage of the total number of steps. The lower the number of slips, the better the slip resistance of the footwear.

6.2 Results

As shown in Figure 23a, the total number of both heel and toe slips are higher for F1 than for F5. The figure also shows that F1 had about twice the number of slips as F5 (33% vs. 19%, respectively) considering all participants and all types of slips. Figure 23b shows that all participants experienced more toe and heel slips with F1 than with F5, except one (Sub4 with 3% vs. 5% for heel slips).

It is worth noting that the large standard deviation in the number of slips ranging from 3 to 20% for heel slips is likely due to different gait patterns among participants. However, the footwear ranking does not change when looking at each participant individually or as an entire group, as opposed to between participants. In other words, F1 was less slip-resistant in most cases than F5.

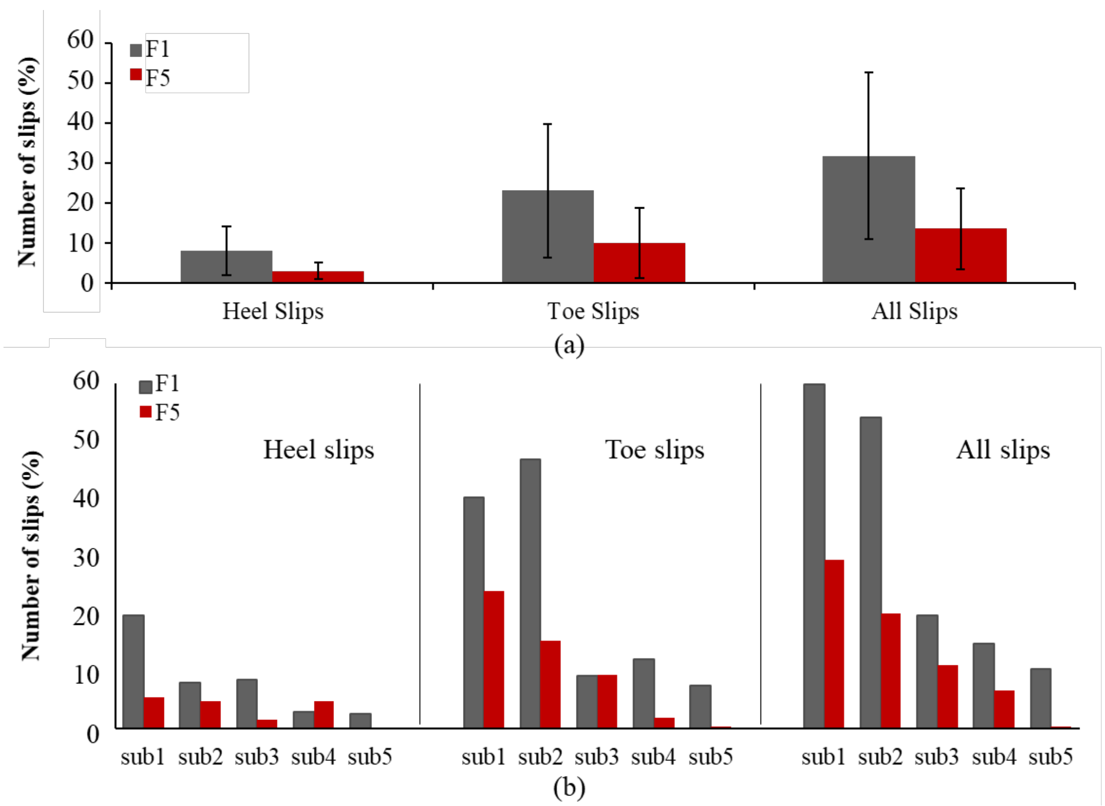
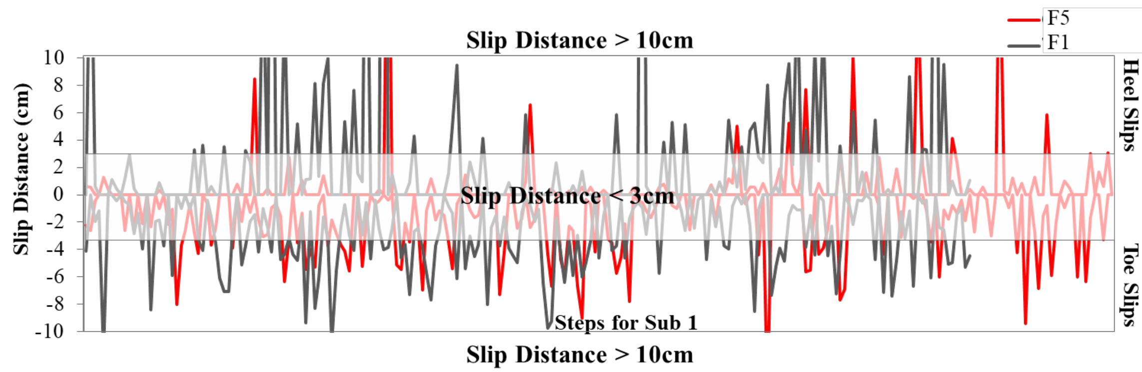


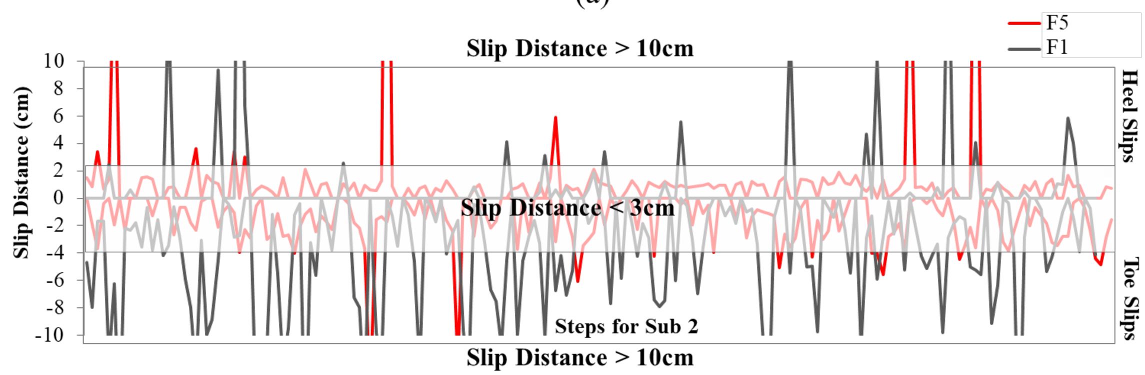
Figure 23. (a) Total number of slips by all participants wearing footwear F1 and F5 while walking on wet ice (the vertical bars represent standard deviations). (b) Number of slips per participant with F1 and F5. Slips are defined as a distance the foot travelled beyond 3 cm after heel contact for heel slips, and beyond 0 cm after toe-off for toe slips. The number of slips was calculated as a percentage of the total number of steps.

The differences in gait patterns are highlighted in Figure 24. The figure shows the region where the slip distance is greater than 3 cm for heel slips and 10 cm for both heel and toe slips, for each participant. For example, Figure 24 (a) and (b) show that both subjects 1 and 2 slipped more than others. In addition, Figure 24 (d) indicates that subject 4 had no hazardous slips (≥ 10 cm) during walking on wet ice in the WinterLab.

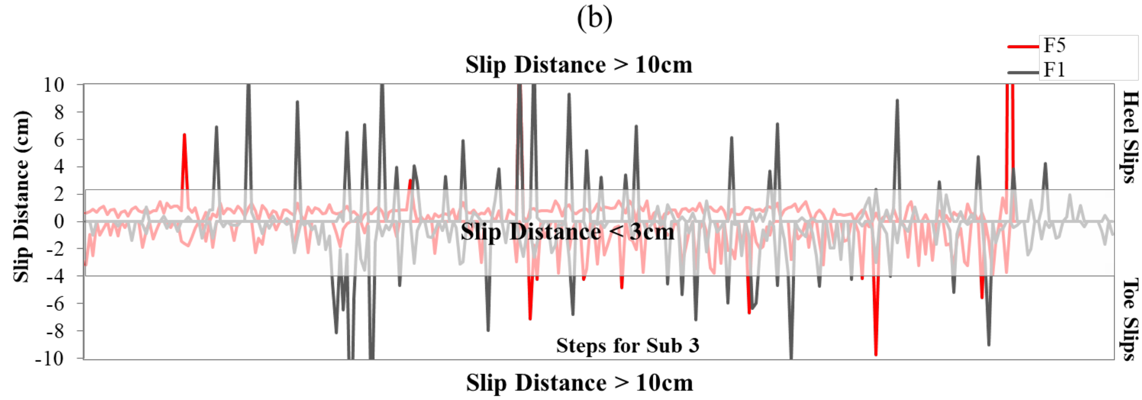
Figure 25 shows that F1 had twice the number of hazardous heel slips than F5 (2.1% vs. 0.9%, respectively). All participants experienced more hazardous slips with F1 than with F5. In addition, all participants responded that F1 was more slippery than F5.


F5
F1
Slip Distance > 10cm
Slip Distance (cm)
Slip Distance < 3cm
Heel Slips
Toe Slips
Steps for Sub 1
Slip Distance > 10cm

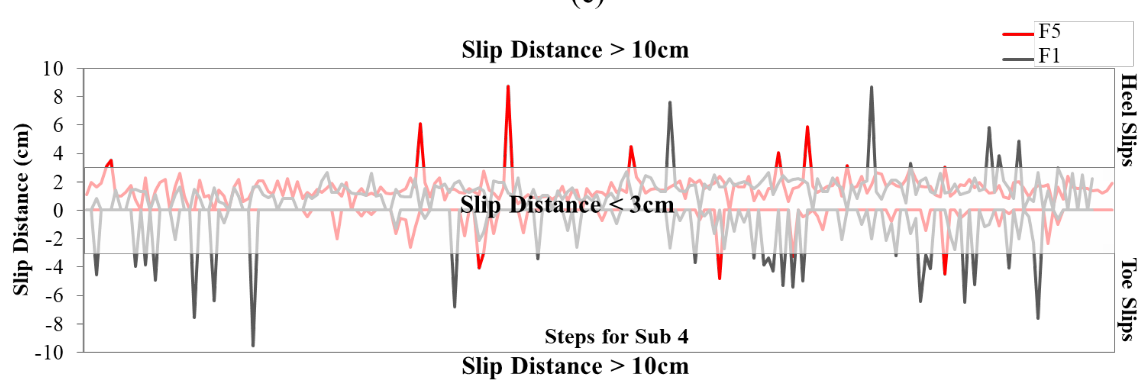
(a)


F5
F1
Slip Distance > 10cm
Slip Distance (cm)
Slip Distance < 3cm
Heel Slips
Toe Slips
Steps for Sub 2
Slip Distance > 10cm

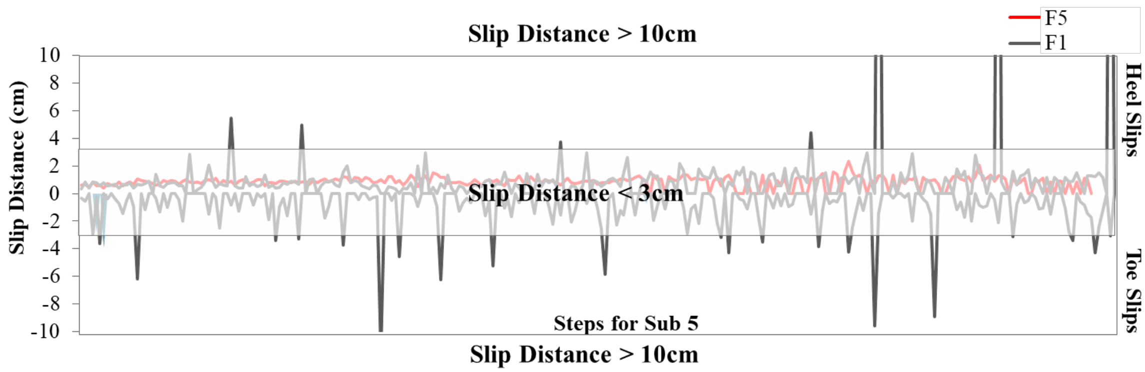
(b)


F5
F1
Slip Distance > 10cm
Slip Distance (cm)
Slip Distance < 3cm
Heel Slips
Toe Slips
Steps for Sub 3
Slip Distance > 10cm

(c)


F5
F1
Slip Distance > 10cm
Slip Distance (cm)
Slip Distance < 3cm
Heel Slips
Toe Slips
Steps for Sub 4
Slip Distance > 10cm

(d)



(e)

Figure 24. Slip distance for both heel and toe slips across all steps taken by (a) subject 1, (b) subject 2, (c) subject 3, (d) subject 4 and (e) subject 5, while walking on wet ice.

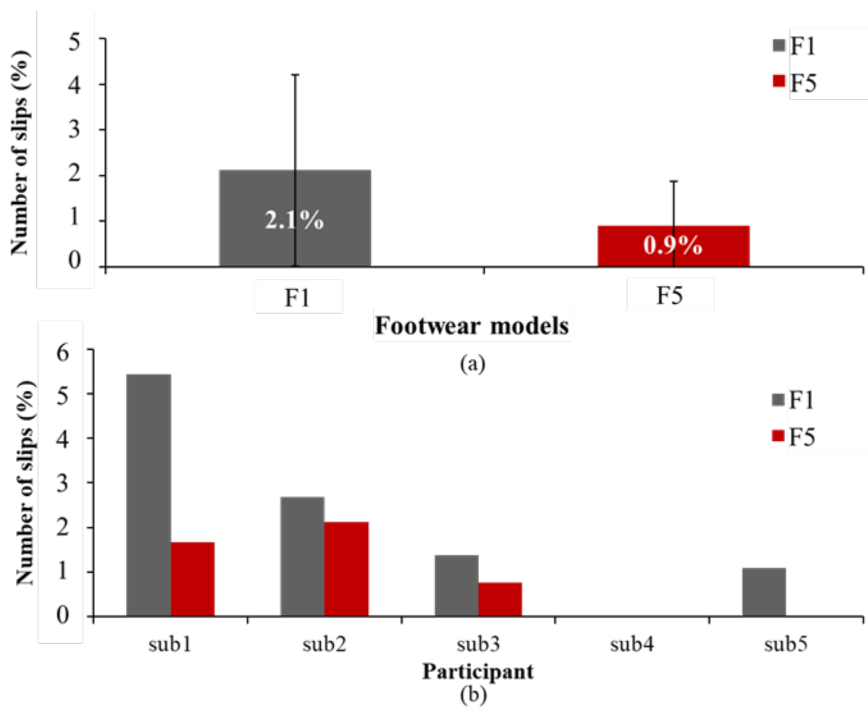


Figure 25. (a) Total number of hazardous heel slips (≥ 10 cm) by all participants wearing footwear F1 and F5 (vertical bars represent standard deviations), (b) Number of hazardous slips per participant with F1 and F5. All data obtained while participants were walking on wet ice.

6.3 Discussion

The comparison between the two boots, based on the four parameters (and on the participant feedback), indicates that F1 is less slip-resistant than F5. This outcome is consistent with the results from the MAA test method, but not with those from the mechanical test method. The fact that the MAA and level walking tests were done under identical environmental conditions (wet ice inside the WinterLab) was certainly a factor in the similarity of the results. It should be noted, however, that even when walking conditions were fairly different (specific walking pace, level walking back and forth instead of ascending and descending), participants experienced the same performances with the two boots.

Even if there was high variability between participants, the phase 3 test protocol, based on the number of slips and on the slip distance (instead of COF), has good potential for footwear comparison. However, the detection of slip events using acceleration data is very challenging. More research will be needed to improve this method and help it become more mature. It may be useful to test other conditions that produced conflicting outcomes, such as on dry ice with boots F1 and F4 (F1 < F4 with the mechanical method, but F1 ≈ F4 with MAA), or boots F8 and F10 (F8 > F10 and F8 = F10 with the mechanical method at IRSST and at KITE, respectively, but F8 << F10 with MAA). Field tests could perhaps help better determine the accuracy of the test methods.

7. GENERAL DISCUSSION

7.1 Advantages and limitations of methods

Test methods like those presented in this study serve to assess footwear sole slip resistance on a variety of surfaces and determine which provide the best performance. When valid, these test methods can be useful in a number of different situations. For instance, to help an occupational health and safety committee select the most appropriate boots for workers, or to help footwear manufacturers develop new, better-performing soles. As the actual performance under real winter conditions is not known, human-centred approaches can generally be considered the gold standard in assessing slip resistance, in comparison with mechanical-based approaches (Grönqvist, Abeysekera, et al., 2001; Grönqvist, Chang, et al., 2001; Hunwin et al, 2010; Powers et al., 2010).

The MAA method provides results that are more externally valid, as it involves having human subjects walk on icy surfaces under controlled conditions. As far as we know, the KITE team is the only one that conducts slip testing in a climate chamber with human subjects. Studies have shown that the MAA method can be a useful, valid test for assessing the performance of boots anywhere (sports and leisure environment, daily activity route as well as workplaces). A study has shown that workers providing personal home support or care, who for eight weeks wore winter boots with a high MAA score, recorded rates of slips and falls that were, respectively, 68% and 78% lower than for workers who wore their own winter boots (Bagheri, Patel, Li, Rizzi, et al., 2019). Another study has also shown that outdoor workers wearing boots with a high MAA score had lower slip rates than outdoor workers wearing their own winter boots (Bagheri et al., 2020). However, the MAA method is not as accessible on a large scale as a mechanical method, although the KITE research team has the capacity and resources to cooperate with other organizations in the development of high-performance soles.

A mechanical method is generally cheaper and more accessible than a human-centred method. That is the attractive aspect of the method developed in this study for the evaluation of COF on ice surfaces. It uses the STM 603 slip resistance tester, already recognized in standardized test methods (ASTM International, 2019b, F2913-19; ISO, 2019, 13287:2019). To the best of our knowledge, no repeatability and reproducibility study has ever been done with this device on ice surfaces. Similarly, it is the first time that work has shown differences between the refrigeration cycles of two SATRA refrigerated ice trays and their possible impact on COF measurement. The method developed in this research study provides acceptable repeatability and reproducibility on a **wet ice** surface. In addition, it has shown good agreement with the MAA method, giving similar results. The MAA method does, however, seem better at taking into account subtle differences between boots, as shown by the test results of boots F1 and F5, with the level walking test method (phase 3). Even if, at first glance, the mechanical method seems to have a good ability to differentiate between boots, it did not find any difference in that case, between F1 and F5 boots.

For **dry ice**, the mechanical method has demonstrated good repeatability and reproducibility within the same laboratory. However, reproducibility between labs has not yet been fully achieved. The agreement and correlation between the two labs are good and the boot ranking is very similar, but the values obtained at IRSST were systematically higher (by around 0.06) than those at KITE. Further research is therefore needed to reduce the bias between the labs. Furthermore, one of the major challenges with the mechanical method on dry ice is how to improve its agreement with the MAA method. Significant differences were noted in the ranking of the boots. Some boots obtained COFs that seemed overestimated in relation to the MAA_COFs, so potentially assessed above their real performance level. For example, if the threshold of 7° with the MAA method (corresponding to a COF of 0.12) developed by KITE is considered, the MAA method would have determined that only boots F9 and F10 among the 10 models tested had a satisfactory performance (Table 12). The mechanical method would have determined that in addition to boots F9 and F10, there were also other boots that had achieved this threshold (six boots, F3 to F8, at IRSST and 3 boots, F4, F7 and F8, at KITE, Table 5).

The mechanical method needs to be improved to increase its ability to estimate boot slip resistance better and come closer to the levels observed with human-centred approaches. For instance, studies have already shown that changes in the test parameters (applied normal force, surface slip speed, heel contact angle) can result in better accuracy of the mechanical method using a whole shoe tester with wet floor tiles (Blanchette & Powers, 2015; Beschorner et al., 2019; Hunwin et al., 2010). In its current state, the mechanical method developed could be used, for instance, to make an initial selection (fast and inexpensive) among different boots, but the MAA method would have to be used for the final selection. Once the reliability of the mechanical method has been improved, it could possibly be adopted as a standardized test method. The small sample size was one of the limitations of this study. Other limitations include a relatively short walking distance for the testing protocol in phases 2 and 3, where only young participants were tested.

7.2 Performance of tested boots

The boots that performed the best were F9 and F10. These boots stood out from the other models, according to the results of the MAA method on both ice surfaces, as well as according to the mechanical method results on both surfaces, but especially on wet ice.

Boot F10 has an Arctic Grip sole (Vibram Corporation, Brookline, MA, USA, us.vibram.com), which includes microscopic fibres embedded in the rubber sole. This boot achieved a mean MAA score of 9° (MAA COF = 0.16) on dry ice and of 14° (0.26) on wet ice. With the mechanical method, this boot also showed the best slip resistance, on both wet and dry ice (at IRSST and KITE). Some of these types of material may lose their slip resistance after wear (Anwer, Bagheri, Fernie, Dutta, & Naguib, 2017), although recent studies have shown improvement in this regard (Bagheri, Anwer, et al., 2019).

Boot F9 has a Green Diamond sole (Mark's/L'Équipeur, Canada), which contains grit embedded into the sole material to create a rough surface. It had a mean MAA score of 9° (0.16) on dry ice and of 10° (0.17) on wet ice. With the mechanical method, this boot also showed excellent slip resistance, i.e., the 2nd best on wet ice (after boot F10) and the 6th best on dry ice (at IRSST and KITE). The mechanical method may have underestimated the performance of this boot, given that the chosen experimental unit (the mean of the last 5 runs) is less suited to soles with grit or crampons (the 1st run may be more appropriate).

Boots F9 and F10 are rated on the web site www.ratemytreads.com in accordance with the rating system developed by KITE. Note that, in contrast with the COF_MAA used in this study, which corresponds to the mean of the MAA scores obtained for the 4 participants, the MAA score used for the snowflake rating corresponds to the lowest MAA measured under the different test conditions (lowest of 16 MAAs, i.e., 4 participants x 2 directions x 2 ice surfaces, dry and wet). The snowflake rating score is a more conservative score to ensure maximum protection against slipping. It should also be noted that the list of boots tested by KITE using the MAA method is regularly updated. Checking out www.ratemytreads.com ratings is an efficient way to select boots that offer good performance.

The study has demonstrated that measuring COFs on different ice surfaces can be useful as a way of getting a more global picture of boot performance. It has shown that COF values on wet ice were generally lower than those on dry ice, meaning that wet ice was more slippery in most cases, regardless of the test method used (mechanical or MAA). The only exception, boot F10, performed better on wet ice than on dry, for both the MAA and mechanical methods at KITE. This rather unusual performance was observed precisely because tests (MAA notably) were conducted on a surface that, *a priori*, could have seemed too slippery for differentiating between soles. It would, therefore, be useful to run tests on at least two types of ice surface, such as the wet and dry ice surfaces used in this study.

It is worth noting that, for the mechanical method, the mean of the three modes of slipping (heel, flat, forepart) was used in the analyses presented. This is the right choice insofar as the purpose was to compare two test methods on the basis of the overall performance ratings they gave to boots, i.e., overall for the 3 modes of slipping for the mechanical method and for the 2 directions of slope for the MAA method. In addition, overall the results showed the same tendencies for all the modes of slipping. At the same time, other choices could be made to assess boot performance with the mechanical method, such as using only the heel mode, as a number of researchers have done (Blanchette & Powers, 2015; Beschorner et al., 2019; Hunwin et al., 2010), as it is the mode of slipping that involves the highest risk of falling, and therefore of serious injury (Grönqvist, Chang, et al., 2001).

7.3 Recommendations

A number of recommendations can be made on the basis of this study's findings.

- Improve the mechanical method in order to achieve satisfactory repeatability and reproducibility, and to better match MAA results, through:
  - A better control of the testing environment (temperature, relative humidity, air velocity inside the lab) and of the testing procedures.
  - An adjustment of the test parameters of the mechanical method (e.g., sliding velocity, normal force, contact angle) to more accurately reflect the MAA results.
- Investigate the use of motion capture data in future data collection with phases 2 and 3 protocols to:
  - Control for walking speed.
  - Include participants from a wider range of ages and abilities.
  - Provide measures of slip distances and develop analysis methods that can take into account the slip distance when comparing boot performance.

- Undertake a sample size calculation and interim power analysis with pilot data collected in phase 3 and recruit a sufficient number of participants to achieve a power of 80%.

- Carry out additional tests using the level walking method (phase 3) on dry ice to confirm which footwear ranking (from the mechanical or the MAA method) is the more reliable.

- To address the limitation of having a relatively short walkway, future work should include a protocol similar to that of phase 3 but taking place outdoors in real-world icy conditions, while the numbers of slip events are tracked by a wearable slip-detector (currently in development).

Because the mechanical method is not reliable enough in its current state, it may be used to make an initial selection from among different boots, but the MAA method would have to be used for the final selection.

Generally speaking, to make the test methods valid and useful, it is necessary to take different ice conditions (e.g., wet and dry ice) into consideration to allow a more detailed evaluation of footwear slip resistance levels, and to assess the performance of footwear in real outdoor environments.

8. CONCLUSION

The goal of this study was to evaluate a mechanical method using the SATRA STM 603 whole shoe tester to determine footwear slip resistance on two different ice surfaces and to compare it with the MAA human-centred test method developed by KITE. The study was conducted in three phases.

The first part of phase 1 involved the development of an alternative mechanical method based on existing test protocols and on determining ice conditions closest to the KITE WinterLab ice conditions. Monitoring of the ice tray's ice temperature with thermistors at the IRSST and at the KITE laboratories revealed that the ice surface temperature is colder than the set point temperature, and it fluctuates as a function of the ice refrigeration cycle. This fluctuation showed slightly different patterns between the two labs. Specific temperature set points and restricted temperature ranges for testing on wet and dry ice surfaces were determined for each lab to ensure the ice temperatures measured by the thermistors were as similar as possible at the two labs and matched the WinterLab's ice temperatures. The set point temperature was 2°C at IRSST and 1°C at KITE for wet ice, and -2°C at both labs for dry ice. The test protocol developed was based on the ASTM F2913-11 standard (ASTM International, 2011). A series of 5 to 10 successive runs was performed for each specific boot in a given experimental condition, and the COF was considered to be the average of the last 5 consecutive runs, generally showing variation of less than 10%. The preparation of the test surface entailed wiping the ice with a wet cloth to remove the frost at the beginning of each series of tests, and ensuring that the tests were run within a specific temperature range (e.g., when the ice tray indicated a temperature between 2°C and 3°C for wet ice at IRSST).

The second part of phase 1 consisted in evaluating the repeatability and reproducibility of this mechanical test method by measuring the slip resistance of 10 types of occupational footwear on wet and dry ice surfaces in different slip modes at the IRSST and KITE laboratories. The intra-laboratory variability was generally acceptable. Regarding inter-laboratory reproducibility, the analyses showed that the measurements taken on **wet ice** were equivalent in both labs in terms of both the COF values and the ranking of the boots, in which boots F9 and F10 stood out as distinctly better performers than the other models. For **dry ice**, the agreement between the two labs was relatively good and the ranking of the boots was more or less equivalent. However, the COF values recorded at IRSST were systematically higher (with a bias of around 0.06) than those obtained at KITE. Greater standardization of the entire test procedure at the two labs, including the ability to control temperature and relative humidity in the labs, would be useful to improve the reproducibility of the mechanical test method.

Phase 2 consisted of evaluating the COFs using the human-centred MAA method and comparing the results with those obtained by the mechanical method. The COFs and boot ranking obtained on **wet ice** with the mechanical method were close to those obtained with the MAA method. Both methods gave distinctly higher COFs for boots F9 and F10 than for the other boots. The correlation between the two methods on wet ice should be interpreted carefully as there was not a substantial COF gradation in the tested boots. The analysis showed good agreement between the two methods, both at IRSST and KITE.

However, on **dry ice**, the mechanical method did not show good agreement with the MAA method, and its boot ranking was quite different. The MAA method found virtually no significant difference between boots F1 to F8 and ranked them as the lowest-performing boots. In contrast, the mechanical method (at both IRSST and KITE) not only found significant differences between many of these boots but also rated boots F4, F7 and F8 as being among the best performers. It may, therefore, be possible that the mechanical method overestimates the performance of some boots in certain cases. It is worth noting, nevertheless, that the mechanical method, like the MAA method, found that boots F9 and F10 offered good slip resistance.

Phase 3 was carried out because phase 2 showed that COF values were not always consistent between the two methods. It consisted of using another human-centred method, based on a motion tracking system, to investigate which method was more reliable for ranking footwear. As there is a greater risk of slipping on wet ice than on dry ice, wet ice was chosen for this part of the study. Boots F1 and F5 were selected since they showed conflicting rating outcomes on wet ice conditions, i.e., the mechanical tests indicated that their COFs were not statistically different, while MAA showed that the two boots had significantly different slip resistance. The comparison between the two boots, based on number of heel slips, number of toe slips, number of hazardous heel slips, and participant feedback, indicated that F1 was less slip resistant than F5, and this outcome was consistent with the results from the MAA test method, but not with those from the mechanical test method.

Regarding the overall performance of the tested boots, the study showed that the boots which performed the best were F9 and F10. They stood out from the other models, according to the results of the MAA method on both ice surfaces, as well as according to the mechanical method results on both surfaces, but especially on wet ice. The study also demonstrated that measuring COFs on different ice surfaces can be useful as a way of getting a more accurate picture of boot performance. It showed that COF values on wet ice were generally lower than those on dry ice, meaning that wet ice was more slippery in most cases, regardless of the test method used (mechanical or MAA).

Recommendations were proposed to improve the mechanical method and to reduce the variability between the labs. One of the major challenges would be to increase the ability of this test method to estimate boot slip resistance better and come closer to the levels observed with human-centred approaches, especially on dry ice. The differences observed between the mechanical and MAA test methods may be due to several testing conditions that were not exactly the same (e.g. ambient air temperature and humidity, footwear conditioning), although efforts were made to obtain ice surface temperatures for the mechanical tests as similar as possible to the WinterLab's ice surface temperatures. They may also be due to the mechanical method's inability to simulate boot dynamic heel-to-toe roll typical of human gait. In its current state, the mechanical method developed could be used, for instance, to do an initial selection (fast and inexpensive) from among different boots, but the MAA method would have to be used for the final selection. Once the reliability of the mechanical method has been improved, it could possibly be adopted as a standardized test method.

BIBLIOGRAPHY

Accessibility for Ontarians with Disabilities Act, S.O. 2005, c. 11.

Anwer, A., Bagheri, Z. S., Fernie, G., Dutta, T., & Naguib, H. E. (2017). Evolution of the coefficient of friction with surface wear for advanced surface textured composites. *Advanced Materials Interfaces, 4*(6). doi: 10.1002/admi.201600983

ASTM International. (2011). *Standard test method for measuring the coefficient of friction for evaluation of slip performance of footwear and test surfaces/flooring using a whole shoe tester*. ASTM standard F2913-11. West Conshohocken, PA: ASTM International.

ASTM International. (2019a). *Interlaboratory study to establish precision statements for ASTM F2913-19, Test method for measuring the coefficient of friction for evaluation of slip performance of footwear and test surfaces/flooring using a whole shoe tester* (Report No. RR:F13-1005). West Conshohocken, PA: ASTM International.

ASTM International. (2019b). *Standard test method for measuring the coefficient of friction for evaluation of slip performance of footwear and test surfaces/flooring using a whole shoe tester*. ASTM standard F2913-19. West Conshohocken, PA: ASTM International.

Bagheri, Z. S., Anwer, A. O, Fernie, G., Naguib, H. E., & Dutta, T. (2019). Effects of multi-functional surface-texturing on the ice friction and abrasion characteristics of hybrid composite materials for footwear. *Wear*, *418-419*, 253-264.

Bagheri, Z. S., Beltran, J., Holyoke, P., Sole, G., Hutchinson K., & Dutta T. (2020). Reducing the risk of falls by 78% with a new generation of slip resistant winter footwear. In R. H. M. Goossens, & A. Murata (Eds.), *Advances in social and occupational ergonomics: Proceedings of the AHFE 2019 International Conference on Social and Occupational Ergonomics, July 24-28, 2019, Washington D. C., USA* (p. 279-285). Cham, Switzerland: Springer.

Bagheri, Z. S., Patel, N., Li, Y., Morrone, K., Fernie, G., & Dutta, T. (2019). Slip resistance and wearability of safety footwear used on icy surfaces for outdoor municipal workers. *Work, 62*(1), 37-47.

Bagheri, Z. S., Patel, N., Li, Y., Rizzi, K., Lui, K. Y. G., Holyoke, P., . . . Dutta, T. (2019). Selecting slip resistant footwear for personal support workers to use outdoors in the winter. *Work, 64*(1), 135-151.

Bentley, T. A., & Haslam, R. A. (1998). Slip, trip and fall accidents occurring during the delivery of mail. *Ergonomics*, *41*(12), 1859-1872.

Beschorner, K. E., Iraqi, A., Redfern, M. S., Cham, R., & Li, Y. (2019). Predicting slips based on the STM 603 whole-footwear tribometer under different coefficient of friction testing conditions. *Ergonomics*, *62*(5), 668-681. doi: 10.1080/00140139.2019.1567828

Blanchette, M. G., & Powers, C. M. (2015). Slip prediction accuracy and bias of the SATRA STM 603 whole shoe tester. *Journal of Testing and Evaluation*, *43*(3), 491-498.

Bland, J. M., & Altman, D. G. (2010). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, *47*(8), 931-936.

Boucher, A. (2019). *Accepted occupational injuries representing cases of fall at the same level or slip-trip without falling, Québec, 2014-2016.* Montreal, QC: IRSST. [Unpublished data].

Cen, D. J. Y. (2018). *Development of a slip analysis algorithm: Automating the maximal achievable angle footwear slip resistance test*. (Master thesis, University of Toronto, Toronto, ON).

Chang, W.-R., Grönqvist, R., Leclercq, L., Myung, R., Makkonen, L., Strandberg, L., . . . Thorpe, S. C. (2001). The role of friction in the measurement of slipperiness, part 1: Friction mechanisms and definition of test conditions. *Ergonomics*, *44*(13), 1217-1232.

Chang, W.-R., Grönqvist, R., Leclercq, L., Brungraber, R. J., Mattke, U., Strandberg,L., . . . Courtney, T. K. (2001). The role of friction in the measurement of slipperiness, part 2: Survey of friction measurement devices. *Ergonomics*, *44*(13), 1233-1261.

Chang, W.-R., Matz, S., Chang, C.-C. (2013). The available coefficient of friction associated with different slip probabilities for level straight walking. *Safety Science*, *58*, 49-52.

Chang, W.-R., Leclercq, S., Lockhart, T. E., & Haslam, R. (2016). State of science: Occupational slips, trips and falls on the same level. *Ergonomics*, *59*(7), 861-883. doi: 10.1080/00140139.2016.1157214

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Courtney, T. K., Sorock, G. S., Manning, D. P., Collins, J. W., & Holbein-Jenny, M. A. (2001). Occupational slip, trip, and fall-related injuries: Can the contribution of slipperiness be isolated? *Ergonomics, 44*(13), 1118-1137.

CSA Group. (2019). *Protective footwear*. CSA Z195-14 (R2019). Mississauga, ON: CSA Group.

Di Pilla, S. (2010). *Slip, trip, and fall prevention: A practical handbook* (2nd ed.). Boca Raton, FL: CRC Press.

Ells, W. (Ed.) (2014). *Safety and occupational footwear*. West Conshohoken, PA: ASTM International.

Gao, C., & Abeysekera, J. (2004). A systems perspective of slip and fall accidents on icy and snowy surfaces. *Ergonomics*, *47*(5), 573-598.

Gao, C., Abeysekera, J., Hirvonen, M., & Grönqvist, R. (2004). Slip resistant properties of footwear on ice. *Ergonomics*, *47*(6), 710-716.

Gao, C., Holmér, I., & Abeysekera, J. (2008). Slips and falls in a cold climate: Underfoot surface, footwear design and worker preferences for preventive measures. *Applied Ergonomics*, *39*(3), 385-391.

Gard, G., & Berggård, G. (2006). Assessment of anti-slip devices from healthy individuals in different ages walking on slippery surfaces. *Applied Ergonomics*, *37*(2), 177-186.

Gauvin, C., Pearsall, D., Damavandi, M., Michaud-Paquette, Y., Farbos, B., & Imbeau, D. (2015). *Risk factors for slip accidents among police officers and school crossing guards* (Report No. R-893). Montreal, QC: IRSST.

Grönqvist, R., Abeysekera, J., Gard, G., Hsiang, S. M., Leamon, T. B., Newman, D. J., . . . Pai, C. Y. (2001). Human-centred approaches in slipperiness measurement. *Ergonomics*, *44*(13), 1167-1199.

Grönqvist, R., Chang, W., Courtney, T. K., Leamon, T . B., Redfern, M. S., & Strandberg, L. (2001). Measurement of slipperiness: Fundamental concepts and definitions. *Ergonomics*, *44*(13), 1102-1117.

Grönqvist, R., & Hirvonen, M. (1995). Slipperiness of footwear and mechanisms of walking friction on icy surfaces. *International Journal of Industrial Ergonomics*, *16*(3), 191-200.

Grönqvist, R. (1995). *A dynamic method for assessing pedestrian slip resistance*. (Report No. 2). Helsinki, Finland: Finnish Institute of Occupational Health.

Hintze, J. (2004). *Number Cruncher Statistical Systems*. Kaysville, UT: NCSS.

Hsu, J. (2015). *Measuring the slip resistance of winter footwear*. (Doctoral thesis, University of Toronto, Toronto, ON).

Hsu, J., Li, Y., Dutta, T., & Fernie, G. (2015). Maximum achievable inclines for footwear while walking on icy slopes and cross-slopes. *Applied Ergonomics*, *50*, 218-225.

Hsu, J., Shaw, R., Novak, A., Li, Y., Ormerod, M., Newton, R., . . . Fernie, G. (2016). Slip resistance of winter footwear on snow and ice measured using maximum achievable incline. *Ergonomics*, *59*(5), 717-728.

Hunwin, G., Ormerod, K., & Darby, A. (2010). A study of the effect of modifying the European standard mechanical slip resistance test for footwear (Report No. RR801). Buxton, England: HSL.

IBM Corporation. (2015). SPSS© Statistics for Windows, version 23.0 [Computer software]. Armonk, NY: IBM Corp.

International Organization for Standardization. (2011). *Personal protective equipment: Safety footwear*. ISO standard 20345:2011. Geneva, Switzerland: ISO.

International Organization for Standardization. (2012a). *Guidance for the selection, use and maintenance of safety and occupational footwear and other personal protective equipment offering foot and leg protection*. ISO/TR standard 18690:2012. Geneva, Switzerland: ISO.

International Organization for Standardization. (2012b). *Personal protective equipment: Occupational footwear*. ISO standard 20347:2012. Geneva, Switzerland: ISO.

International Organization for Standardization. (2014). *Personal protective equipment: Protective footwear*. ISO standard 20346:2014. Geneva, Switzerland: ISO.

International Organization for Standardization. (2019). *Personal protective equipment: Footwear: Test method for slip resistance*. ISO standard 13287:2019. Geneva, Switzerland: ISO.

James, D. I. (1999). The theory behind the DIN ramp tests. *Polymer Testing*, *18*(1), 3-10.

Jung, K., & Fischer, A. (1993). An ISO test method for determining slip resistance of footwear: Determination of its precision. *Safety Science*, *16*(2), 115-127.

Leamon, T. B., & Li, K. W. (1990). *Microslip length and the perception of slipping*. Paper presented at the 23rd International Congress on Occupational Health, Montreal, QC (p. 17).

Loo-Morrey, M. (2006). Ramp testing natural and man made stone floors (Report No. HSL/2006/47). Buxton, England: HSL. Retrieved from http://www.hse.gov.uk/research/hsl_pdf/2006/hsl0647.pdf

Manning, D. P., Ayers, I., Jones, C., Bruce, M., & Cohen, K. (1988). The incidence of underfoot accidents during 1985 in a working population of 10,000 Merseyside people. *Journal of Occupational Accidents, 10*(2), 121-130.

NCSS Statistical Software. (2016). NCSS 11 [Computer software]. Kaysville, UT: NCSS.

Ormerod, K. (2010). *Monitoring the slip resistance of emerging workplace footwear* (Report No. RR780). Buxton, England: HSL.

Powers, C. M., Blanchette, M. G., Brault, J. R., Flynn, J., & Siegmund, G. P. (2010). Validation of walkway tribometers: Establishing a reference standard. *Journal of Forensic Sciences*, *55*(2), 366-370.

SATRA Technology. (2011). *Friction (slip resistance) of footwear and floorings* (Method No. TM144:2011). Kettering, England: SATRA Technology Centre.

Strandberg, L. (1985). The effect of conditions underfoot on falling and overexertion accidents. *Ergonomics*, *28*(1), 131-147.

Strandberg, L., & Lanshammar, H. (1981). The dynamics of slipping accidents. *Journal of Occupational Accidents. 3*(3), 153-162.

Swedler, D. I., Verma, S. K., Huang, Y. H., Lombardi, D. A., Chang, W. R., Brennan, M., & Courtney, T. K. (2015). A structural equation modelling approach examining the pathways between safety climate, behaviour performance and workplace slipping. *Occupational & Environmental Medicine*, *72*(7), 476-481.

Tisserand, M. (1985). Progress in the prevention of falls caused by slipping. *Ergonomics*, *28*(7), 1027-1042.

Wilson, M. (1990). Development of SATRA slip test and tread pattern design guidelines. In B. Gray (Ed.), *Slips, stumbles, and falls: Pedestrian footwear and surfaces* (pp. 113-123). West Conshohocken, PA: ASTM International.

Wilson, M. (1996). Slip resistance characteristics of footwear solings assessed using the SATRA friction tester. *Journal of Testing and Evaluation*, *24*(6), 377-385.

APPENDIX A: ADDITIONAL RESULTS FOR THE DEVELOPMENT OF THE MECHANICAL TEST METHOD

Preliminary tests were conducted with the STM 603 slip resistance tester to (1) determine the set point temperatures of the ice machine for making wet and dry ice surfaces similar to the targeted ones of the WinterLab, (2) develop the test protocol to achieve satisfactory repeatability and reproducibility and (3) decide on the basic experimental unit to use to estimate the COF for a given condition.

A.I Determination of ice protocol

On the basis of the ice surface temperature readings with the Mid_top thermistor (Figure 4), the ice machine set point temperatures were identified that had the best potential to produce ice surfaces similar to those of the WinterLab. The results (Table 14) show that for dry ice conditions, the set point temperature of the ice machine that produces a mean ice surface temperature closest to that of the WinterLab (-5.55°C) is -2°C for the KITE lab (which gives a mean ice surface temperature of -5.50°C). For the IRSST lab, the set point temperature of -2°C leads to a mean ice surface temperature of approximately 1°C less cold (i.e., -4.55°C) than that of the WinterLab, but that set point temperature was chosen nevertheless for dry ice because the difference between the minimum and maximum values of the cycle (range = 4.19°C) is less than when the set point temperature is colder (at set point = -3°C, range = 5.46°C).

Table 14. Ice surface temperature according to various ice tray temperature set points in the IRSST and KITE labs, and ice surface temperature in the WinterLab

Ice	Lab	SATRA ice machine	Ice surface temperature measured by thermistor Mid_top (°C)					
		Set point (°C)	Max	Mean	Min	Range	SD	Cycle time (min)
Dry	WinterLab	-	-5.09	-5.55	-6.07	0.98	0.31	27.0
	KITE	-1	-2.39	-4.87	-6.33	3.94	1.25	7.7
		-2	**-3.06**	**-5.50**	**-7.36**	**4.30**	**1.05**	**11.3**
	IRSST	**-2**	**-2.52**	**-4.55**	**-6.71**	**4.19**	**1.15**	**6.4**
		-3	-2.22	-4.94	-7.68	5.46	1.54	5.6
Wet	WinterLab	-	-0.90	-1.57	-2.09	1.19	0.38	31.0
	KITE	2	0.76	-1.32	-3.79	4.55	1.41	7.3
		1	**0.41**	**-1.76**	**-3.68**	**4.09**	**1.25**	**6.7**
		0	-0.47	-2.84	-4.41	3.94	1.13	7.7
	IRSST	**2**	**-0.33**	**-2.47**	**-4.11**	**3.78**	**1.16**	**6.2**
		1	-0.88	-3.25	-5.31	4.43	1.28	8.9

For wet ice conditions (Table 14), the set point temperature of the ice machine that produces a mean ice surface temperature closest to that of the WinterLab (-1.57°C) is -1°C for the KITE lab (ice temperature = -1.76°C), with a difference of just 0.19°C. For the IRSST lab, the set point temperature of 1°C leads to a mean ice surface temperature of around 1.68°C colder (i.e., -3.25°C) than that of the WinterLab. In this case, the set point temperature of 2°C was chosen for the IRSST lab. It gives a mean ice temperature of -2.47°C, i.e., a difference of 0.90°C. It is worth noting that, at these set point temperatures, the coldest part of the ice tray is permanently covered with a thin layer of water, which almost never freezes. But that area is fairly small and lies outside the area where the boots are tested.

Table 14 also shows that the temperature of the ice of the WinterLab varies far less than the ice surface temperatures of the ice machines, in particular because the temperature of the air in the WinterLab is colder than the air of the IRSST and KITE labs, where the ice machines are located.

For the selected set point temperatures, a test window was determined with a view to identifying at what point in the cycle the temperatures are as close as possible to the WinterLab temperatures (shaded areas of Figure 5). During these test windows, the mean temperatures at the ice surface were fairly close to that of the WinterLab, with differences of less than 1°C for dry ice and less than 1.2°C for wet ice, as shown in Table 15.

Table 15. Ice surface temperatures at selected ice tray temperature set points for dry and wet ice conditions at the IRSST and KITE labs

Ice	Lab	SATRA ice machine	Ice surface temperature measured by thermistor Mid_top (°C)					
		Set point (°C)	Max	Mean	Min	Range	SD	Testing time slot (min)
Dry	WinterLab	-	-5.09	-5.55	-6.07	0.98	0.31	27.0
	KITE	-2	-5.78	-6.28	-6.88	-1.10	0.39	3.90
	IRSST	-2	-3.63	-5.12	-6.71	3.08	0.80	3.70
Wet	WinterLab	-	-0.90	-1.57	-2.09	1.19	0.38	31.0
	KITE	1	0.23	-0.41	-1.11	-1.34	0.58	1.00
	IRSST	2	-1.11	-1.86	-2.6	1.49	0.43	1.87

A.II Differences between newly developed mechanical testing protocol and existing test methods

Table 16 summarizes the main differences between the newly developed mechanical testing protocol and the existing test methods ASTM F2913-11 and SATRA TM144:2011. All other test parameters were as specified in the ASTM F2913-11 standard.

Table 16. Summary of protocol comparison between standard and developed test methods

	ASTM F2913-11 on indoor surfaces	SATRA TM144:2011 guidelines on ice surfaces	Newly developed mechanical test method KITE-IRSST with SATRA STM 603 machine on ice surfaces
Footwear conditioning	23°C 50%RH	-7°C in cooling bath for 2 hours	21–23°C 29–45%RH
Testing surfaces	• Quarry tiles or stainless steel • Dry or contaminated (water, soap, glycerin)	Ice tray at -7°C: • Frosted ice • Smooth dry ice • Resurface the ice using the hot knife	• Ice tray at: ○ -2°C dry ice ○ 2°C wet ice (1°C for KITE) • Frost removed by wiping using a wet cloth (used also to resurface the ice)
COF	COF = Mean COF of last 5 runs	Frosted ice → COF = 1st run Dry smooth ice → COF = 4th run	COF = Mean COF of last 5 runs

A.III Footwear characterization

The composition of the sole materials was not known. The hardness of the soles (Shore A) was measured in the IRSST lab at a temperature of 24°C (the values are given in Table 1). The sole hardness was also measured when cold. The 10 models of boot were placed in an incubator (Forma™ environmental chamber model 3851, Thermo Fisher Scientific Inc., Waltham, MA, USA) at various temperatures for 3 hours (10°C, 0°C and -7°C, constant relative humidity of 50%) and the hardness of the soles was measured immediately upon removing them from the incubator. The results are shown in Figure 26.

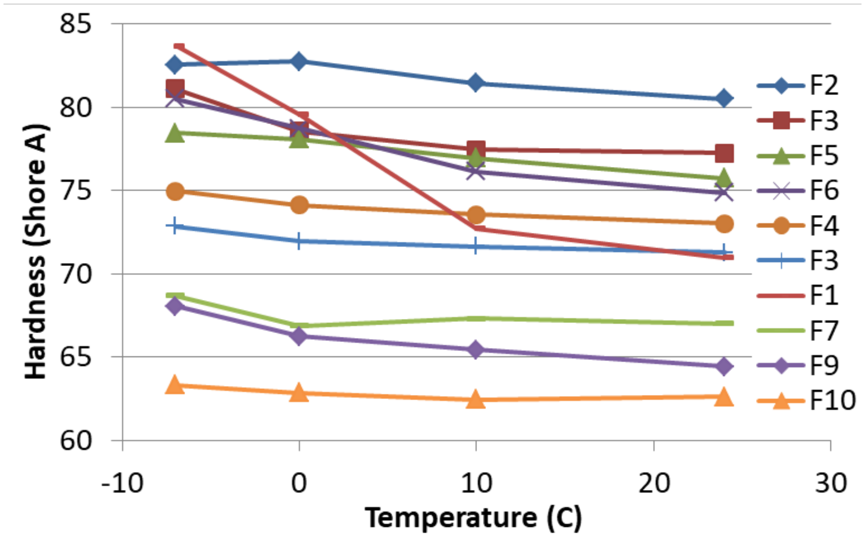


Figure 26. Sole hardness as a function of temperature, for the 10 footwear types.

Overall, the hardness of the soles remained constant between 0°C and 23°C, except for boot F1, where the sole was harder at 0°C. This result was expected, given that the boots being tested were winter boots and that their soles should normally retain their flexibility in cold.

For one boot, the temperature of the sole was measured as a function of time using an infrared reader (model Fluke 62 MAX+, Fluke Corp., Everett WA, USA), from the time it was removed from the incubator at -7°C until it reached the ambient temperature. These measurements indicated that in the first 10 minutes, the sole warmed up by 1°C for each minute, approximately.

The results showed that conditioning the boots at the laboratory temperature could be satisfactory for the mechanical method.

A.IV A *posteriori* validation of experimental unit

Because the SlipMASTER software does not automatically calculate the coefficient of variation (CV) of the last 5 runs, and because some boots showed greater variation, the coefficient of variation of the last 5 runs was estimated *a posteriori*, that is, after the completion of phase 1 for the 540 conditions tested, to check whether or not it was indeed 10% or less in most cases.

The results (Table 17) showed that the CVs were 10% or less for virtually all the measurements done on dry ice (99% of cases, i.e., 268/270 measurements). However, two thirds (67%) of the 270 measurements taken on wet ice had CVs of 10% or less. For wet ice, the COFs estimated at the IRSST lab seemed a little more stable during the series from 5 to 10 runs, as 84% of the 180 measurements taken at that lab had CVs of 10% or less, compared with 34% of the 90 measurements taken at the KITE lab. Variability was generally greater for wet ice than for dry ice, given that the COF measurements were lower and very close to the resolution of the data acquisition system of the STM 603 device (0.01).

Table 17. Percentage of cases where the coefficient of variation of the last 5 consecutive runs for the COF measurement series is less than or equal to 10%

Ice	Lab	Mode	Nb of cases CV ≤ 10%	Total nb of cases*	%	%/ice/lab	%/ice	%		%/lab
DRY	IRSST	Flat	60	60	100	100%	99%	83%	IRSST	92%
		Forepart	60	60	100					
		Heel	60	60	100					
	KITE	Flat	29	30	97	98%				
		Forepart	29	30	97					
		Heel	30	30	100					
WET	IRSST	Flat	47	60	78	84%	67%		KITE	66%
		Forepart	55	60	92					
		Heel	49	60	82					
	KITE	Flat	12	30	40	34%				
		Forepart	8	30	27					
		Heel	11	30	37					

* Total number of cases corresponds to 10 types of footwear x 3 repetitions x 1 or 2 operators.

APPENDIX B: DETAILED MECHANICAL AND MAA TEST RESULTS (PHASES 1 AND 2)

B.I COFs with mechanical test method at the IRSST and KITE labs in flat, forepart and heel slip modes

Table 18. COFs of 10 types of footwear tested on dry ice in three slip modes by two operators at the IRSST lab and by one operator at the KITE lab. CVs are ≤ 10% in 73% of cases, and ≤ 15% in 93% of cases

Footwear	IRSST (operator 1, 3 repetitions)									IRSST (operator 2, 3 repetitions)									KITE (one operator, 3 repetitions)								
	Flat			Forepart			Heel			Flat			Forepart			Heel			Flat			Forepart			Heel		
	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV
F1	0.134	0.005	3.9%	0.083	0.012	13.9%	0.095	0.005	5.3%	0.128	0.011	8.7%	0.099	0.001	1.2%	0.102	0.002	2.0%	0.096	0.010	10.8%	0.079	0.002	2.9%	0.053	0.015	28.6%
F2	0.112	0.003	3.1%	0.103	0.008	8.1%	0.115	0.001	1.0%	0.134	0.007	5.2%	0.114	0.005	4.6%	0.133	0.011	8.3%	0.054	0.009	16.1%	0.051	0.008	15.7%	0.040	0.010	25.0%
F3	0.144	0.007	4.8%	0.110	0.010	9.1%	0.144	0.016	11.4%	0.176	0.022	12.3%	0.111	0.010	9.1%	0.165	0.016	9.7%	0.100	0.020	20.0%	0.070	0.000	0.0%	0.083	0.009	11.3%
F4	0.257	0.015	6.0%	0.195	0.020	10.4%	0.221	0.023	10.4%	0.247	0.034	13.6%	0.191	0.007	3.7%	0.236	0.014	5.9%	0.172	0.007	4.2%	0.139	0.001	0.8%	0.133	0.006	4.8%
F5	0.184	0.017	9.5%	0.136	0.023	17.0%	0.159	0.020	12.4%	0.177	0.009	5.3%	0.123	0.010	8.0%	0.167	0.012	7.2%	0.145	0.016	10.7%	0.075	0.005	6.7%	0.093	0.007	7.5%
F6	0.150	0.000	0.0%	0.137	0.015	11.2%	0.171	0.009	5.3%	0.140	0.010	7.1%	0.133	0.015	11.5%	0.174	0.022	12.8%	0.106	0.002	1.9%	0.057	0.006	10.2%	0.094	0.004	4.3%
F7	0.259	0.020	7.8%	0.197	0.017	8.6%	0.235	0.005	2.1%	0.219	0.010	4.7%	0.193	0.022	11.4%	0.228	0.018	8.0%	0.201	0.016	8.1%	0.143	0.013	8.9%	0.134	0.002	1.5%
F8	0.321	0.011	3.4%	0.229	0.003	1.3%	0.236	0.012	5.1%	0.325	0.016	4.9%	0.253	0.008	3.0%	0.297	0.040	13.5%	0.229	0.011	4.8%	0.171	0.008	4.7%	0.185	0.008	4.4%
F9	0.202	0.011	5.5%	0.183	0.005	2.8%	0.191	0.012	6.0%	0.180	0.005	2.9%	0.182	0.005	2.9%	0.193	0.022	11.2%	0.159	0.012	7.7%	0.118	0.010	8.8%	0.123	0.010	8.3%
F10	0.315	0.036	11.4%	0.197	0.006	2.9%	0.190	0.010	5.3%	0.266	0.023	8.5%	0.195	0.005	2.6%	0.193	0.005	2.4%	0.245	0.005	2.1%	0.203	0.014	7.0%	0.158	0.011	6.7%
Pooled SD	0.016			0.014			0.013			0.017			0.011			0.019			0.012			0.008			0.009		
Pooled SD	0.014									0.016									0.010								
Pooled SD	0.014																										

Table 19. COFs of 10 types of footwear tested on wet ice in three slip modes by two operators at the IRSST lab and one operator at the KITE lab. CVs are ≤ 10% in 36% of cases, and ≤ 15% in 64% of cases

Footwear	IRSST (operator 1, 3 repetitions)									IRSST (operator 2, 3 repetitions)									KITE (one operator, 3 repetitions)								
	Flat			Forepart			Heel			Flat			Forepart			Heel			Flat			Forepart			Heel		
	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV
F1	0.053	0.006	10.8%	0.046	0.005	11.5%	0.020	0.000	0.0%	0.061	0.012	18.8%	0.043	0.006	13.3%	0.053	0.008	15.2%	0.059	0.005	8.6%	0.043	0.008	18.9%	0.017	0.004	25.0%
F2	0.042	0.002	4.8%	0.030	0.010	33.3%	0.029	0.010	35.1%	0.073	0.012	15.7%	0.040	0.010	25.0%	0.044	0.005	12.0%	0.031	0.002	7.4%	0.031	0.006	20.5%	0.027	0.006	22.4%
F3	0.057	0.008	14.5%	0.035	0.003	8.8%	0.039	0.009	22.9%	0.077	0.011	14.7%	0.039	0.001	2.9%	0.059	0.012	20.0%	0.057	0.005	8.8%	0.031	0.005	14.7%	0.030	0.005	17.6%
F4	0.077	0.010	12.8%	0.051	0.001	2.3%	0.041	0.008	19.6%	0.075	0.009	12.7%	0.055	0.009	16.7%	0.075	0.008	11.2%	0.087	0.004	4.8%	0.075	0.009	12.5%	0.068	0.010	15.3%
F5	0.039	0.008	20.5%	0.053	0.006	12.2%	0.033	0.005	14.1%	0.037	0.011	29.5%	0.035	0.013	36.4%	0.020	0.000	0.0%	0.061	0.013	21.0%	0.011	0.005	47.2%	0.013	0.005	36.5%
F6	0.055	0.015	27.5%	0.063	0.006	9.1%	0.050	0.000	0.0%	0.066	0.007	10.5%	0.069	0.019	27.4%	0.037	0.006	15.7%	0.049	0.002	4.7%	0.019	0.008	40.6%	0.017	0.002	13.3%
F7	0.071	0.009	12.8%	0.059	0.009	15.2%	0.047	0.005	9.8%	0.065	0.010	15.7%	0.073	0.015	20.8%	0.040	0.010	25.0%	0.087	0.013	14.9%	0.047	0.006	13.8%	0.039	0.004	10.8%
F8	0.109	0.002	2.1%	0.083	0.021	25.0%	0.071	0.012	16.4%	0.105	0.005	4.8%	0.066	0.007	10.5%	0.099	0.013	13.1%	0.099	0.008	8.2%	0.081	0.008	10.3%	0.081	0.009	11.7%
F9	0.181	0.015	8.3%	0.181	0.011	6.1%	0.194	0.021	10.8%	0.189	0.010	5.3%	0.181	0.001	0.6%	0.180	0.002	1.1%	0.173	0.007	4.1%	0.135	0.017	12.7%	0.135	0.002	1.7%
F10	0.233	0.007	3.0%	0.209	0.017	8.0%	0.181	0.020	11.1%	0.219	0.010	4.6%	0.203	0.004	2.1%	0.156	0.007	4.4%	0.291	0.009	3.1%	0.221	0.020	9.1%	0.165	0.006	3.9%
Pooled SD	0.009			0.011			0.011			0.010			0.010			0.008			0.008			0.011			0.006		
Pooled SD	0.010									0.009									0.008								
Pooled SD	0.009																										

Table 20. COFs of 10 types of footwear tested on dry and wet ice conditions in three slip modes by the two operators at the IRSST lab. For dry ice, CVs are ≤ 10% in 70% of cases, and ≤ 15% in 97% of cases. For wet ice, CVs are ≤ 10% in 23% of cases, and ≤ 15% in 40% of cases

Footwear	Dry ice									Wet ice								
	IRSST (6 repetitions)									IRSST (6 repetitions)								
	Flat			Forepart			Heel			Flat			Forepart			Heel		
	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV
F1	0.131	0.008	6.5%	0.091	0.011	12.5%	0.098	0.005	5.4%	0.057	0.009	16.2%	0.045	0.005	11.6%	0.037	0.019	51.7%
F2	0.123	0.013	10.6%	0.109	0.009	7.9%	0.124	0.012	9.5%	0.058	0.019	32.4%	0.035	0.010	30.0%	0.036	0.011	30.4%
F3	0.160	0.023	14.2%	0.110	0.009	8.1%	0.154	0.018	11.9%	0.067	0.014	21.0%	0.037	0.003	8.9%	0.049	0.014	28.8%
F4	0.252	0.024	9.5%	0.193	0.014	7.1%	0.229	0.019	8.2%	0.076	0.009	11.5%	0.053	0.006	12.1%	0.058	0.020	33.9%
F5	0.181	0.013	7.2%	0.130	0.017	13.4%	0.163	0.015	9.3%	0.038	0.009	22.7%	0.044	0.013	29.9%	0.026	0.008	28.6%
F6	0.145	0.008	5.8%	0.135	0.014	10.2%	0.172	0.015	8.9%	0.060	0.012	20.2%	0.066	0.013	19.6%	0.043	0.008	18.8%
F7	0.239	0.026	11.0%	0.195	0.018	9.1%	0.231	0.013	5.4%	0.068	0.009	13.4%	0.066	0.014	20.5%	0.044	0.008	18.4%
F8	0.323	0.013	3.9%	0.241	0.014	5.9%	0.267	0.043	16.0%	0.107	0.004	3.8%	0.075	0.017	22.5%	0.085	0.019	22.2%
F9	0.191	0.014	7.5%	0.182	0.005	2.5%	0.192	0.016	8.1%	0.185	0.012	6.5%	0.181	0.007	3.9%	0.187	0.015	8.2%
F10	0.291	0.038	13.1%	0.196	0.005	2.5%	0.191	0.007	3.7%	0.226	0.011	4.8%	0.206	0.011	5.5%	0.168	0.019	11.3%
Pooled SD	0.020			0.012			0.019			0.011			0.011			0.015		
Pooled SD	0.018									0.013								

B.II MAA COFs in downhill and uphill slip directions

Table 21. Downhill and uphill MAA COFs of 10 types of footwear (mean of 4 participants, SD, and coefficient of variation) tested on dry and wet ice conditions

Foot wear	Dry – Downhill			Dry – Uphill			Wet – Downhill			Wet – Uphill		
	COF	SD	CV	COF	SD	CV	COF	SD	CV	COF	SD	CV
F1	0.046	0.015	31.5%	0.039	0.019	49.2%	0.026	0.016	61.7%	0.026	0.016	61.7%
F2	0.043	0.017	39.0%	0.030	0.007	22.5%	0.046	0.016	34.9%	0.046	0.016	34.9%
F3	0.051	0.017	33.4%	0.040	0.015	38.9%	0.066	0.015	23.6%	0.066	0.015	23.6%
F4	0.062	0.016	26.5%	0.077	0.014	18.7%	0.066	0.035	52.8%	0.066	0.035	52.8%
F5	0.036	0.012	34.0%	0.028	0.025	88.8%	0.070	0.028	40.0%	0.070	0.028	40.0%
F6	0.057	0.014	25.4%	0.028	0.016	57.1%	0.072	0.017	24.0%	0.072	0.017	24.0%
F7	0.059	0.015	25.5%	0.058	0.024	41.6%	0.072	0.015	20.2%	0.072	0.015	20.2%
F8	0.089	0.020	22.5%	0.087	0.014	16.1%	0.074	0.022	30.1%	0.074	0.022	30.1%
F9	0.184	0.013	6.8%	0.148	0.023	15.6%	0.174	0.025	14.6%	0.174	0.025	14.6%
F10	0.200	0.028	13.8%	0.226	0.054	24.0%	0.261	0.009	3.5%	0.261	0.009	3.5%

B.III Overall COFs obtained with mechanical and MAA test methods

Table 22. Overall COFs and MAA COFs of 10 types of footwear tested on dry ice

Footwear	IRSST Overall COF			KITE Overall COF			MAA COF		
	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV
F1	0.107	0.020	18.3%	0.076	0.020	26.2%	0.090	0.029	32.0%
F2	0.119	0.013	10.8%	0.048	0.010	20.6%	0.081	0.021	25.6%
F3	0.142	0.028	19.9%	0.084	0.016	19.5%	0.112	0.021	18.6%
F4	0.224	0.031	13.8%	0.148	0.018	12.4%	0.072	0.024	32.8%
F5	0.158	0.026	16.5%	0.104	0.031	30.0%	0.087	0.023	26.1%
F6	0.151	0.020	13.4%	0.086	0.022	25.6%	0.112	0.016	14.3%
F7	0.222	0.027	12.2%	0.159	0.032	20.1%	0.087	0.030	33.7%
F8	0.277	0.044	15.8%	0.195	0.027	13.6%	0.094	0.016	17.0%
F9	0.189	0.013	6.7%	0.134	0.021	16.0%	0.158	0.016	10.2%
F10	0.226	0.052	22.9%	0.202	0.037	18.5%	0.158	0.013	8.3%

Table 23. Overall COFs and MAA COFs of 10 types of footwear tested on wet ice

Footwear	IRSST Overall COF			KITE Overall COF			MAA COF		
	Mean	SD	CV	Mean	SD	CV	Mean	SD	CV
F1	0.046	0.015	31.5%	0.039	0.019	49.2%	0.026	0.016	61.7%
F2	0.043	0.017	39.0%	0.030	0.007	22.5%	0.046	0.016	34.9%
F3	0.051	0.017	33.4%	0.040	0.015	38.9%	0.066	0.015	23.6%
F4	0.062	0.016	26.5%	0.077	0.014	18.7%	0.066	0.035	52.8%
F5	0.036	0.012	34.0%	0.028	0.025	88.8%	0.070	0.028	40.0%
F6	0.057	0.014	25.4%	0.028	0.016	57.1%	0.072	0.017	24.0%
F7	0.059	0.015	25.5%	0.058	0.024	41.6%	0.072	0.015	20.2%
F8	0.089	0.020	22.5%	0.087	0.014	16.1%	0.074	0.022	30.1%
F9	0.184	0.013	6.8%	0.148	0.023	15.6%	0.174	0.025	14.6%
F10	0.200	0.028	13.8%	0.226	0.054	24.0%	0.261	0.009	3.5%

APPENDIX C: DETAILED ANALYSIS RESULTS FOR COMPARISON OF MECHANICAL TEST METHOD BETWEEN IRSST AND KITE (PHASE 1)

C.I Global ANOVAs

Table 24. ANOVAs for global mechanical test results, including effect of footwear, mode, ice and lab (significant with $p \leq 0.05$)

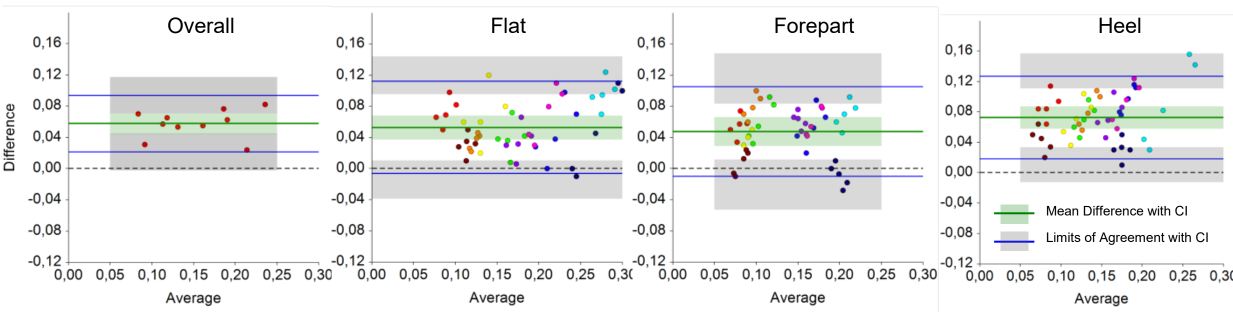
Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
Footwear	1.122	9	.125	469.197	<.001	.897
Ice	.640	1	.640	2408.886	<.001	.832
Mode	.107	2	.053	200.560	<.001	.453
Lab	.125	1	.125	470.056	<.001	.492
Footwear * Ice	.362	9	.040	151.252	<.001	.737
Footwear * Mode	.054	18	.003	11.216	<.001	.294
Footwear * Lab	.023	9	.003	9.411	<.001	.149
Ice * Mode	.018	2	.009	33.547	<.001	.122
Ice * Lab	.078	1	.078	291.789	<.001	.376
Mode * Lab	.008	2	.004	15.332	<.001	.059
Error	.129	485	2.658E-4			
Total (corrected)	3.018	539				

Table 25. ANOVAs for mechanical test results of each ice surface (dry and wet), including effect of footwear, mode and lab (significant with $p \leq 0.05$)

Ice	Source	Type III Sum of Squares	df	Mean Square	F	*p*-value	Partial Eta Squared η_p^2
Dry	Footwear	.601	9	.067	246.320	<.001	.907
	Mode	.078	2	.039	144.433	<.001	.559
	Lab	.200	1	.200	736.876	<.001	.764
	Footwear * Mode	.046	18	.003	9.517	<.001	.429
	Mode * Lab	.007	2	.003	12.745	<.001	.101
	Footwear * Lab	.018	9	.002	7.577	<.001	.230
	Error	.062	228	2.710E-4			
	Total (corrected)	1.112	269				
Wet	Footwear	.812	9	.090	540.484	<.001	.955
	Mode	.040	2	.020	119.267	<.001	.511
	Lab	.003	1	.003	16.840	<.001	.069
	Footwear * Mode	.019	18	.001	6.421	<.001	.336
	Mode * Lab	.004	2	.002	13.229	<.001	.104
	Footwear * Lab	.018	9	.002	11.901	<.001	.320
	Error	.038	228	1.670E-4			
	Total (corrected)	1.009	269				

C.II Bland-Altman analyses comparing the two labs using the mechanical test method

a) Dry ice: IRSST vs. KITE



b) Wet ice: IRSST vs. KITE

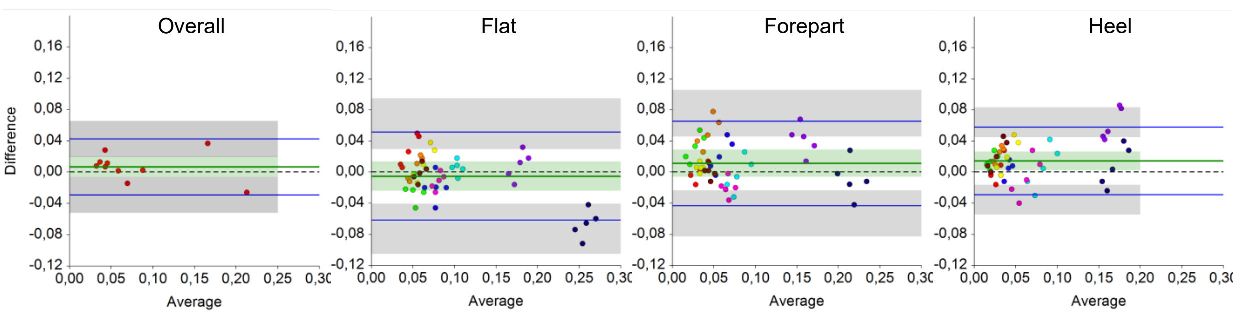


Figure 27. Bland-Altman plots for COF comparison between the IRSST and KITE labs for (a) dry ice and (b) wet ice. The plots[5] are for COFs tested in all three modes (overall), and tested in flat, forepart, and heel modes.

---

[5] For the "overall" plots: The values of $\overline{COF_f}$, combining all the modes of slipping, were used. In order to consider the grouping of all measurements, the Design 1 comparison was chosen in NCSS 11. This design considers that each of the two measurement methods is evaluated once on each boot.

For the "flat", "forepart" and "heel" plots: The values of $\overline{COF_{f,m}}$ were used. In order to take into account the repetitions that were done for each of the conditions, Design 2 was chosen in NCSS 11. This design considers that each boot is measured several times using one method and then measured several times using the other method, without natural pairing of the measures. Since the measurements are not paired, it is somewhat arbitrary which values are plotted. The NCSS software chose to plot five points per footwear, no matter how many possible combinations of points were possible. The five points were generated as follows (ref: NCSS 11 documentation):

i. Determine the minimum, maximum, and average for method 1 (e.g., measurements at KITE).
ii. Determine the minimum, maximum, and average for method 2 (e.g., measurements at IRSST).
iii. Compute four pairs using all combinations of the minimum and maximum values of each method.
iv. Compute one additional pair using the average of method 1 and then the average of method 2.
v. Compute the difference and average values of these five pairs and plot them.

[6] See footnote 5 on page 76, in reference to Figure 27.

Table 26. Bland-Altman analyses for COF comparison between the IRSST and KITE labs for dry and wet ice conditions

	DRY							WET						
	Bias			Limit of Agreement			R	Bias			Limit of Agreement			R
	Value	SD	CI	Value	SD	CI		Value	SD	CI	Value	SD	CI	
Overall	**0.058**	**0.018**	**0.013**	**0.036**	**0.010**	**0.023**	**0.94**	**0.007**	**0.018**	**0.013**	**0.036**	**0.010**	**0.023**	**0.96**
Flat	0.053	0.022	0.016	0.059	0.010	0.025	0.95	-0.005	0.026	0.019	0.057	0.014	0.032	0.96
Forepart	0.048	0.026	0.019	0.057	0.014	0.032	0.87	0.011	0.024	0.018	0.054	0.013	0.030	0.93
Heel	0.073	0.020	0.015	0.054	0.010	0.023	0.92	0.014	0.017	0.012	0.044	0.008	0.019	0.96

Limit of Agreement: Lower value = Bias – Value of LoA
Upper value = Bias + Value of LoA

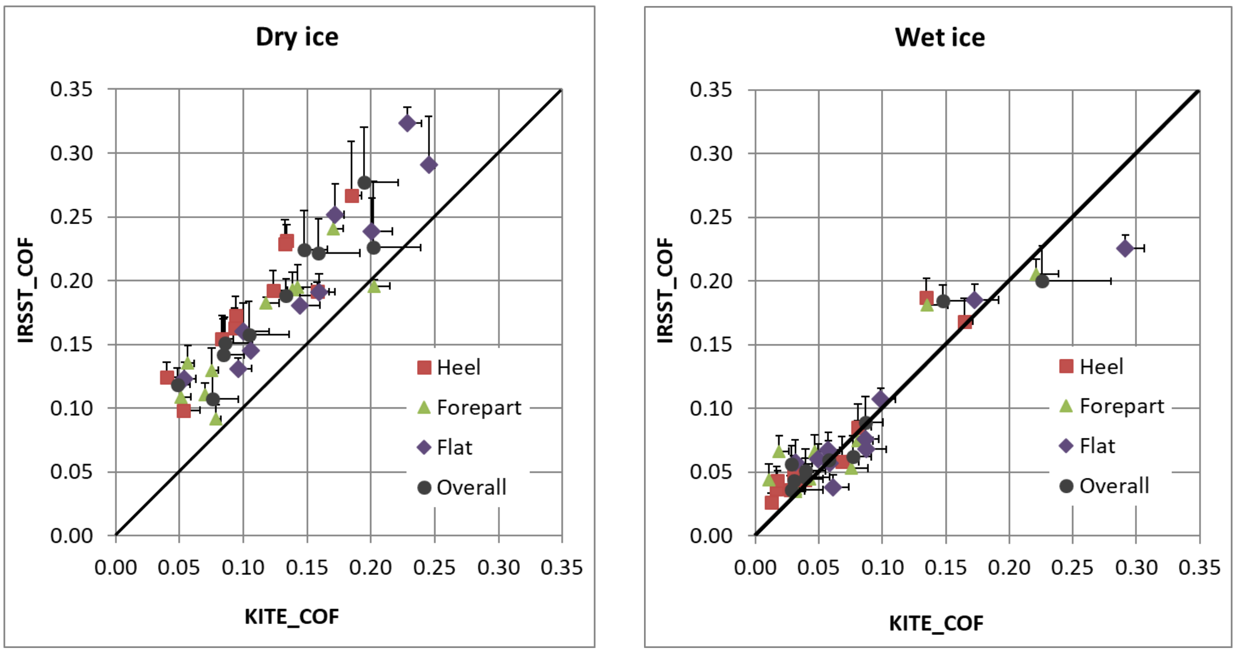


Figure 28. Comparison of IRSST vs. KITE COF for flat, forepart, heel and overall results for 10 types of footwear in dry and wet ice conditions. In the graphs, the black solid line represents what would be a perfect agreement and correlation between IRSST and KITE results.

Table 27. Footwear ranking and significant subsets according to COFs obtained with mechanical test method in flat, forepart and heel modes on dry ice for the two labs (IRSST and KITE)

Foot wear	N	Subsets for **IRSST** – COF **Flat**				
		1	2	3	4	5
F2	6	0.123				
F1	6	0.131				
F6	6	0.145	0.145			
F3	6	0.160	0.160	0.160		
F5	6		0.181	0.181		
F9	6			0.191		
F7	6				0.239	
F4	6				0.252	
F10	6					0.291
F8	6					0.323

Foot wear	N	Subsets for **KITE** – COF **Flat**					
		1	2	3	4	5	6
F2	3	0.054					
F1	3		0.096				
F3	3		0.100				
F6	3		0.106				
F5	3			0.145			
F9	3			0.159			
F4	3			0.172	0.172		
F7	3				0.201	0.201	
F8	3					0.229	0.229
F10	3						0.245

Foot wear	N	Subsets for **IRSST** – COF **Forepart**				
		1	2	3	4	5
F1	6	0.091				
F2	6	0.109	0.109			
F3	6	0.110	0.110			
F5	6		0.130	0.130		
F6	6			0.135		
F9	6				0.182	
F4	6				0.193	
F7	6				0.195	
F10	6				0.196	
F8	6					0.241

Foot wear	N	Subsets for **KITE** – COF **Forepart**					
		1	2	3	4	5	6
F2	3	0.051					
F6	3	0.057	0.057				
F3	3	0.070	0.070				
F5	3		0.075				
F1	3		0.079				
F9	3			0.118			
F4	3			0.139	0.139		
F7	3				0.143		
F8	3					0.171	
F10	3						0.203

Foot wear	N	Subsets for **IRSST** – COF **Heel**					
		1	2	3	4	5	6
F1	6	0.098					
F2	6	0.124	0.124				
F3	6		0.154	0.154			
F5	6			0.163	0.163		
F6	6			0.172	0.172		
F10	6				0.191		
F9	6				0.192		
F4	6					0.229	
F7	6					0.231	0.231
F8	6						0.267

Foot wear	N	Subsets for **KITE** – COF **Heel**				
		1	2	3	4	5
F2	3	0.040				
F1	3	0.053				
F3	3		0.083			
F5	3		0.093			
F6	3		0.094			
F9	3			0.123		
F4	3			0.133	0.133	
F7	3			0.134	0.134	
F10	3				0.158	
F8	3					0.185

Table 28. Footwear ranking and significant subsets according to COFs obtained with mechanical test method in flat, forepart and heel modes on wet ice for the two labs (IRSST and KITE)

Foot wear	N	Subsets for **IRSST** – COF **Flat**				
		1	2	3	4	5
F5	6	0.038				
F1	6	0.057	0.057			
F2	6	0.058	0.058			
F6	6		0.060			
F3	6		0.067			
F7	6		0.068			
F4	6		0.076			
F8	6			0.107		
F9	6				0.185	
F10	6					0.226

Foot wear	N	Subsets for **KITE** – COF **Flat**				
		1	2	3	4	5
F2	3	0.031				
F6	3	0.049	0.049			
F3	3		0.057			
F1	3		0.059			
F5	3		0.061			
F4	3			0.087		
F7	3			0.087		
F8	3			0.099		
F9	3				0.173	
F10	3					0.291

Foot wear	N	Subsets for **IRSST** – COF **Forepart**				
		1	2	3	4	5
F2	6	0.035				
F3	6	0.037				
F5	6	0.044				
F1	6	0.045				
F4	6	0.053	0.053			
F7	6		0.066	0.066		
F6	6		0.066	0.066		
F8	6			0.075		
F9	6				0.181	
F10	6					0.206

Foot wear	N	Subsets for **KITE** – COF **Forepart**					
		1	2	3	4	5	6
F5	3	0.011					
F6	3	0.019	0.019				
F2	3	0.031	0.031				
F3	3	0.031	0.031				
F1	3		0.043				
F7	3		0.047	0.047			
F4	3			0.075	0.075		
F8	3				0.081		
F9	3					0.135	
F10	3						0.221

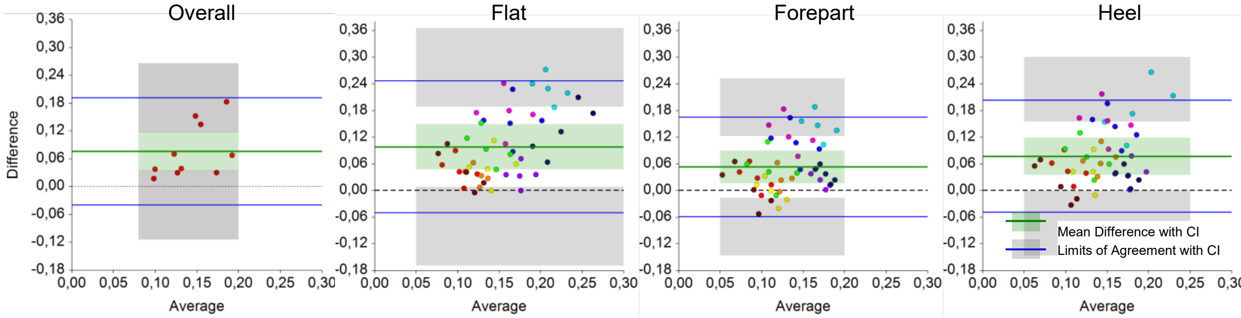
Foot wear	N	Subsets for **IRSST** – COF **Heel**			
		1	2	3	4
F5	6	0.026			
F2	6	0.036	0.036		
F1	6	0.037	0.037		
F6	6	0.043	0.043		
F7	6	0.044	0.044		
F3	6	0.049	0.049		
F4	6		0.058	0.058	
F8	6			0.085	
F10	6				0.168
F9	6				0.187

Foot wear	N	Subsets for **KITE** – COF **Heel**				
		1	2	3	4	5
F5	3	0.013				
F1	3	0.017				
F6	3	0.017				
F2	3	0.027	0.027			
F3	3	0.030	0.030			
F7	3		0.039			
F4	3			0.068		
F8	3			0.081		
F9	3				0.135	
F10	3					0.165

APPENDIX D: DETAILED ANALYSIS RESULTS FOR COMPARISON BETWEEN MECHANICAL AND MAA TEST METHODS (PHASE 2)

D.I Bland-Altman analyses comparing mechanical and MAA test methods

a) Dry ice: IRSST vs. MAA



b) Dry ice: KITE vs. MAA

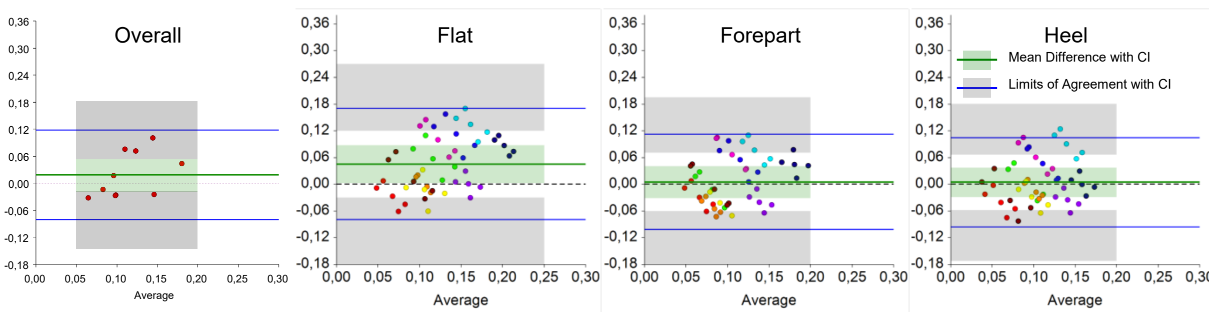
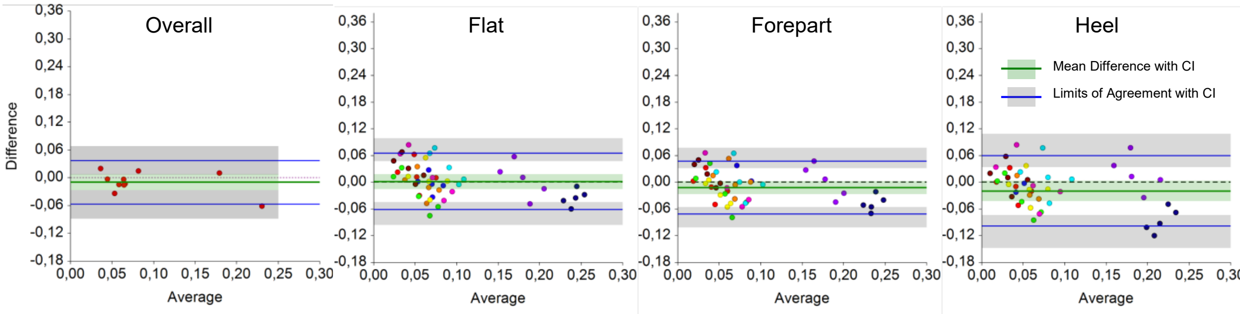


Figure 29. Bland-Altman plots for comparison between mechanical and MAA test results on dry ice for (a) IRSST and (b) KITE. The plots[6] are for COFs tested in all three modes (overall), and tested in flat, forepart, and heel modes.

[6] See footnote 5 on page 76, in reference to Figure 27.

a) Wet ice: IRSST vs. MAA



b) Wet ice: KITE vs. MAA

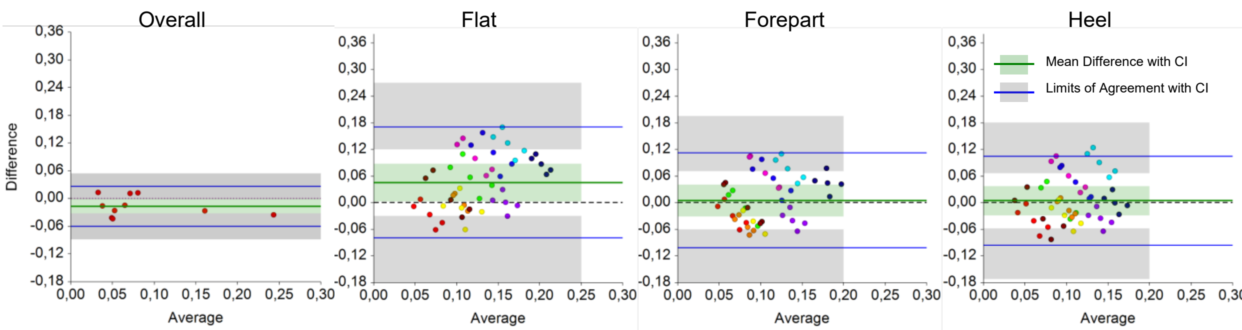


Figure 30. Bland-Altman plots for comparison between mechanical and MAA test methods on dry ice for (a) IRSST and (b) KITE. The plots[7] are for COFs tested in all three modes (overall), and tested in flat, forepart, and heel modes.

[7] See footnote 5 on page 76, in reference to Figure 27.

Table 29. Bland-Altman analyses for comparison between mechanical and MAA test methods on dry and wet ice conditions: (a) IRSST vs. MAA, and (b) KITE vs. MAA.

a) IRSST vs. MAA

	DRY							WET						
	Bias			Limit of Agreement			R	Bias			Limit of Agreement			R
	Value	SD	CI	Value	SD	CI		Value	SD	CI	Value	SD	CI	
Overall	**0.076**	**0.059**	**0.042**	**0.115**	**0.033**	**0.075**	**0.14**	**-0.010**	**0.024**	**0.017**	**0.047**	**0.014**	**0.031**	**0.94**
Flat	0.098	0.071	0.051	0.149	0.038	0.089	0.17	0.002	0.023	0.017	0.063	0.011	0.026	0.95
Forepart	0.053	0.052	0.037	0.112	0.028	0.065	0.21	-0.012	0.021	0.015	0.060	0.010	0.023	0.96
Heel	0.077	0.059	0.042	0.126	0.031	0.073	0.03	-0.019	0.032	0.023	0.079	0.016	0.037	0.89

b) KITE vs. MAA

	DRY							WET						
	Bias			Limit of Agreement			R	Bias			Limit of Agreement			R
	Value	SD	CI	Value	SD	CI		Value	SD	CI	Value	SD	CI	
Overall	**0.018**	**0.051**	**0.036**	**0.099**	**0.028**	**0.064**	**0.34**	**-0.017**	**0.022**	**0.016**	**0.043**	**0.012**	**0.028**	**0.95**
Flat	0.045	0.060	0.043	0.125	0.032	0.075	0.33	0.007	0.020	0.014	0.057	0.009	0.022	0.97
Forepart	0.005	0.050	0.036	0.107	0.027	0.062	0.35	-0.023	0.027	0.019	0.068	0.013	0.031	0.92
Heel	0.004	0.046	0.033	0.100	0.024	0.057	0.31	-0.034	0.031	0.022	0.073	0.016	0.037	0.91