1. Introduction

Wintertime, the number of accidents increases considerably (Brown & Baass 1997), particularly when the roads are covered by snow or ice (Usman et al. 2010). The ability to detect such conditions is therefore desired. This information can be used in order to warn road users using variable message signs (Räma & Kulmala 2000)and adjust the speed limits to the current conditions (Rämä 1999). Knowledge about the current road conditions can also be used by winter maintenance personnel to take suitable measures where and when it is required. There exists several types of sensors to detect current road conditions, such as embedded capacitance sensors (Meindl et al. 2006), visual light cameras combined with image analysis (Kuehnle & Burghout 1998) and radar (Viikari et al. 2009).

Another method, to remotely identify different road surface conditions, is to use the difference in reflectivity of different road conditions (i.e. dry, wet, icy) in the the near infrared (NIR) part of the electromagnetic spectrum (Casselgren et al. 2007; Jonsson et al. 2015). This method has the advantage of a large spatial coverage, the ability to detect areas and not only points, and the ability to be installed as mobile sensors. Today, several commercial sensors using NIR/SWIR absorption are available on the market. What properties the sensors can measure varies depending on the producer, but typically, they can detect the current road condition, estimate a friction value and provide a measure of water film thickness. These parameters are, however, not measured directly but a result of an algorithm based on the reflectance at the different wavelengths which are used. How good a sensor is, will hence depend on its algorithms. When using road condition sensors, it is important that the sensors are reliable (give correct information) and that they are consistent (such that different sensors does not give very different information). Therefore, validation of the performance of such sensors is required in order to better understand what they can be used for and what limitations the sensors have.

Validation of optical sensors in the field is difficult. The non-visible wavelengths used by most sensors makes it difficult to know exactly where a sensor is looking. Since the spatial variations on a road can be large (Kinosita et al. 1970), it becomes difficult to know exactly what a sensor is looking at. While laboratory test may be less realistic, they have the advantage that the surface condition and the location of the observed area can be controlled with good precision. Laboratory studies can hence serve as a “best-case” scenario for the sensors, giving an indication what they are capable of and how to best use the data they prove. In this study, we tested five different sensors in a laboratory environment using different “road conditions”. The study is not a performance test, where the sensors are ranked according to their results. Instead, we focus on how reliable the results of optical road condition sensors are, and how consistent they are relative to each other. This is knowledge we believe can important in order to understand the information optical road condition sensors provides, and to learn about their limitations.

2. Method

2.1 Sensors and sensor setup

In this study five different sensors, two stationary and three mobile (Table 1), were tested in a walk-in cold laboratory with an adjustable temperature. The tested sensors were Vaisala DSC211 (stationary), Metsense 2DRoad (stationary), Teconer RCM411 (mobile), Metsense MetRoad Mobile (mobile) and MARWIS (mobile). At the time of the study, the software versions of XXXXXXXX were used.

The sensors were mounted with the aid of the corresponding producers, in order to get a fair test of sensors in an environment they are not designed to perform in. This was particularly important for the fixed sensors, which could not be mounted at the distance they usually operate at.

The parameters that we compared (Table 1) were classified road condition, derived friction and water film thickness. In addition, ice film thickness, ice percentage and snow-water equivalence were given by some of the sensors, but since each of these only was provided by one sensor, they could not be compared between sensors. Since the mobile sensors can be expected to operate over several different types of asphalt, those were tested on two different types of asphalt. One gray "old" asphalt and one black "new" asphalt. The stationary sensors were only tested on the gray asphalt only. Before the test began, all sensors were calibrated to the gray asphalt plate. The exception was the MetRoad Mobile sensor, which is calibrated using a sheet of white paper.

Table 1 The tested sensors, and the road condition related parameters they provide.

Sensor type	Sensor name	Road condition	Estimated friction	Waterfilm thickness	Ice film thickness	Ice percentage	Snow-water equivalence
Fixed sensors	Vaisala DSC211	X	X	X	X		X
	Metsense 2Droad	X					
Mobile sensors	Teconer RCM411	X	X	X			
	Mensense Metroad-Mobile	X	X				
	Lufft MARWIS	X	X	X		X	

The testing of different road conditions was done by adding different types of contaminants on an asphalt substrate, 0.3 x 0.3 m in area and 50 mm thick. As previously mentioned, two different substrates were made, one with gray and one with black asphalt. Since the sensors operates in approximately the same wavelengths, they could not look at the same spot simultaneously without interfering with each other. This was solved by mounting the sensors such that each looked at a different spot, and the asphalt substrates were moved to the sensor which was up for testing. The substrate placement was fixed relative to each sensor, such that they looked at the same area of the substrate each measurement. The sensing areas for each sensor is shown in Figure 1.

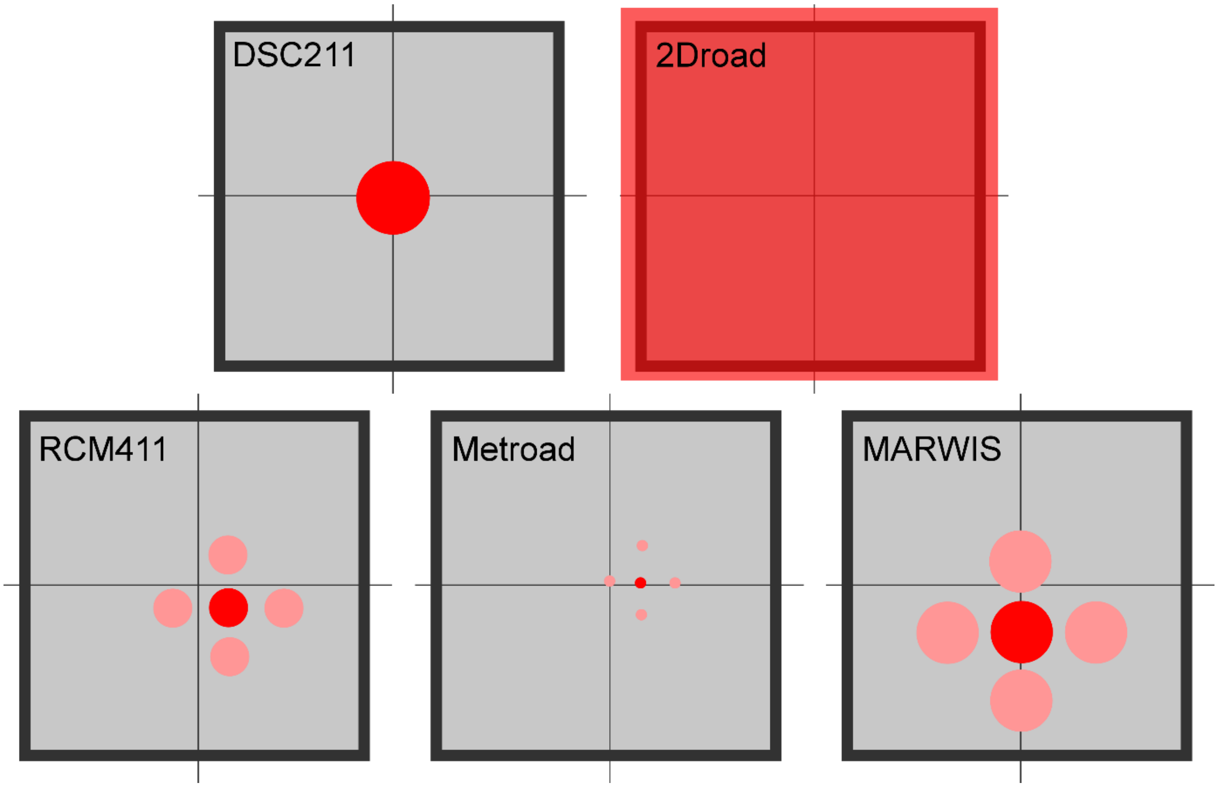

Figure 1 Sensing spots for the different sensors. For the mobile sensors, the red circle indicates the central spot, while the pink circles show the other four positions where measurements were made.

2.2 Surface preparation

Three different types of contaminants were tested in this study; water, ice and snow. For water and ice, the film thickness was varied. For the mobile sensors, we also changed the underlying asphalt type. For snow we only used the gray asphalt, but we used different snow types with different densities to simulate the large range of road conditions which can occur when a road is contaminated by snow (Kinosita). Below follows a description on how the different road conditions were prepared.

Water

Measurements on wet substrates were performed inside the walk-in freezer at a temperature of +10°C. Since asphalt to some extent is permeable, a water film would not be stable over the time required for all sensors to measure it. To solve this, the substrates were submerged in a water bath and the water level was adjusted such that the substrate was covered with five different film thicknesses; 0.5, 1, 2, and 3 mm. The black asphalt substrate was, however, too hydrophobic to allow the thinnest film and thus only water-films from 1-3 mm were prepared on this substrate.

Ice

Measurements of ice-covered substrates were performed at a temperature of -3°C. Water was distributed across the substrates in thin layers using a sponge. The water was allowed to freeze between each subsequent layer. Four different thicknesses were prepared, and the exact thickness was measured using a depth micrometer. The thickness could not be controlled exactly using this method, which gave rise to some slight thickness differences between the substrates (Table 2).

Snow

Snow was produced in a snow-machine built on the principles of the one by (Schleef et al. 2014). This gave dendritic snow (fresh snow), similar to what can be expected to fall on a road. Snow, however, can come in many different types which behaves very differently on a road (Kinosita et al. 1970). In order to test how the sensors handled different types of snow, one third of the fresh snow was allowed to age in a freezer at -3°C for one month while another third was submerged in a 0°C water bath for 48 hours. This way two other snow types were achieved, the first old granular snow and the last large grained "spring snow".

Only the old-asphalt substrate was used together with snow, and for each of the snow types it was covered with a 50 mm thick snow layer. The surface of the snow was levelled using a ruler. This produced snow densities of 150 kg/m^3, 420 kg/m^3 and 530 kg/m^3 for fresh, old and spring-snow respectively. Snow on roads is rarely untouched, however. Traffic and weather compresses and densifies the snow to a large range of densities. This was simulated by compressing the fresh and old snow (the spring snow could not be compressed as the large grains only began to flow off the substrate) using an MTS uniaxial compression rig. The fresh snow was compressed to densities of 450 $kg/m^3$and 750 kg/m^3, while the old snow was compressed to a density of 620 kg/m^3.

To put all these snowtypes and densities into context, the 150 kg/m^3 fresh snow corresponds to dry untouched snow lying on a road. The 450 fresh snow and both old snow types would correspond to the surface of a typical winter road with compacted snow. The spring snow could be found on a winter road towards the end of the season, or on a salted road where the salt rapidly transforms fresh snow into this type of snow consisting of large round grains. Finally, the fresh snow compressed to a density of 750 kg/m^3 (ice has a density of 917 kg/m^3) became extremely hard and very smooth. In all aspects (apart from the color), this surface had more in common with ice than snow.

Winter roads covered with compacted snow are often treated by adding abrasives, such as gravel, to increase friction. This was simulated by adding different amounts of gravel to one substrate covered by compacted snow. The amount of gravel was quantified by the fractional area of the snowy substrate which it covered. Two different gravel amounts were tested, 5% and 25%. The very small footprint of the MetRoad Mobile (about the same size as the gravel), made this sensor unsuitable for the study and hence it was not tested on gravel.

2.3 Measurements

The substrates where placed at fixed positions relative to each sensor. For the mobile sensors, measurements were performed at five different areas on the substrates (Figure 1), in order to accommodate for them being designed to measure over a distance rather than a point measurement.

Each sensor had its own software system to collect data, and raw measurements were averaged over different time scales. The road condition measurements therefore had to be adjusted individually for each sensor. When sensors provided several different classes for the same road condition/coating/contaminant, the predominant class was registered as the result. In cases where a second class was reported in more than 30% of the measurements, this class was also registered.

Vaisala DSC211

The Vaisala DSC211 sensor collected data once per minute. However, the reported class and film thickness was an average over a longer period of time (5-10 min). To accommodate for this time-averaging, the prepared substrate was placed in front of the sensor for 20 minutes before measurements commenced. The DSC211 was then allowed to measure during 30 minutes, giving 30 datapoints per road condition.

Metsense 2D-road

Collected approximately one hyperspectral image every minute. In total, 10 such images were collected for each individual road condition. Each image consists of 256x256 pixels, but to reduce the amount of data and enable a comparison with the other sensors, only the predominating class from these pixels was recorded (unless a secondary class consisted of more than 30% of the surface).

Teconer RCM411

The RCM collected data at 1 Hz, and data was collected for one minute at each of the five measurement areas. This resulted in a total of 300 datapoints per road condition

Metsense MetRoad Mobile

Measured at 50 Hz. 100 datapoints were collected at each of the five measurements areas, giving a total of 500 datapoints per road condition.

Lufft MARWIS

The MARWIS sensor collected raw-data at 100Hz, but averaged this over 15 s when writing the data to an accessible file. The MARWIS collected data over 2 minutes for each of the five measurement areas, giving a total of 40 datapoints per road condition.

3. Results and discussion

3.1 Road condition classification

Table 2 The different contaminants and the corresponding classification results

Surface description			Classification result				
Condition	Plate	Details	DSC211	2Droad	RCM411	Metroad	Marwis
Dry	Gray		Dry	Dry	Dry	Dry	Dry
	Black		-	-	Moist	Dry	Dry
Wet	Gray	0.5 mm	Wet	Moist	Wet	Moist	Wet
		1.0 mm	Wet	Wet	Moist	Frost	Wet
		2.0 mm	Wet	Wet	Slush	Ice	Wet
		3.0 mm	Wet	Wet	Wet	Wet	Wet
	Black	1.0 mm	-	-	Moist	Moist	Wet
		2.0 mm	-	-	Wet + slush	Wet	Wet
		3.0 mm	-	-	Wet	Ice	Wet
Ice	Gray	0.5 mm	Ice	Ice	Ice	Frost	Ice
		0.9 mm	Ice	Ice	Ice + snow	Ice	Ice
		2.2 mm	Ice	Ice	Ice	Ice	Wet
		3.5 mm	Ice	Ice	Ice	Ice	Wet
	Black	0.6 mm	-	-	Ice	Frost	Ice
		0.9 mm	-	-	Ice	Ice	Ice
		2.6 mm	-	-	Ice + slush	Wet	Wet
		3.7 mm	-	-	Ice	Wet	Wet
Snow	Gray	Fresh 150	Snow	Snow	Snow	Snow	Snow
		Fresh 450	Snow	Snow	Snow	Snow	Snow
		Fresh 750	Snow	Snow	Snow	Snow	Snow
		Old 420	Snow	Snow	Snow	Snow	Snow
		Old 620	Snow	Snow	Snow	Snow	Snow
		Spring 530	Ice	Snow	Snow	Snow	Ice + snow
Snow w/ gravel	Gray	0%	Snow	Snow	Snow	-	Snow
		5%	Snow + ice	Snow + ice	Snow + Ice	-	Snow
		25%	Snow	Snow + ice	Ice + snow	-	Snow

The classification results, presented in Table 1, show that the two stationary sensors performed very well in detecting what contaminant which covered the substrate. The only misclassifications of the stationary sensors was that the DSC211 considered coarse spring snow as ice and that both had problems with gravel covered snow. For the mobile sensors, the classification results were more varying. Metroad and MARWIS correctly classified both substrates when dry, while RCM411 classified the black asphalt as being moist. Looking at substrates with water films ranging from 0.5 to 3 mm, two of the sensors (Metroad and RCM411) classified some of the water films as ice and slush respectively. This error persisted also on the black asphalt substrate. The algorithm used by MARWIS, however, correctly classified all wet samples on both plates. When looking at ice, the RCM411 correctly classified all samples on both substrates. Metroad gave correct classifications on the gray substrate, while thick ice (>2 mm) on the black substrate was classified as "wet". MARWIS classified thick ice (>2 mm) as wet on both substrates. The issue the mobile sensors has with distinguishing between water and ice does not seem to be related to the underlying substrate, but rather related to the water/ice film thickness. While a water film thickness of more than 1 mm probably is uncommon on a road, ice films this thick can certainly be expected. An inability to detect such ice films can be serious. This is, after all, one of the more important tasks when detecting road conditions. Both in order to warn road users against slippery conditions which might be hard to detect by eye, and when it comes to inform maintenance personnel that action is required to increase traction on a road.

When the gray substrate was covered with the different snow types and densities, all sensors classified the contaminant as snow. The only exception was the coarse grained spring-snow, which was classified as ice or ice and snow by the DSC211 and the MARWIS respectively. The reason for this is likely the reduced reflection from coarse-grained snow fooling the algorithms that they were looking at ice. This particular snow-type is probably not so common on a road, making this a minor issue. A larger problem with snow, however, is related to how it is defined. Snow exists in many different shapes and densities. On a road, this can result in snow that is blown off the road after the first pass of a car, or snow which is extremely hard and dense and which in every aspect but the color should be considered as ice. In this study, the fresh snow 150 kg/m^3 corresponds to the first case while the fresh snow 750 kg/m^3 corresponds to the second case. Being aware that both these conditions would give the same road condition status from optical sensors is hence important when using this type of sensors during wintertime.

Snow with added gravel was problematic for the sensors. Three of the four tested sensors classified the gravel covered snow as snow and ice. The reason for this is, according to one of the producers, that gravel is not a defined class in the sensor software. When the algorithms sees the gravel, it is classified as the most resembling of the existing classes, which in this case happened to be ice. The only sensor which did not classify the gravel as ice, was the MARWIS. This did, however, only classify the surface as pure snow.

3.2 Friction values

The analysis of the friction coefficients provided by all sensors but the 2Droad, was somewhat different from the classification results. Friction values would have to be validated in the field, and there was hence no correct answer to compare with. However, the friction values acquired in this study can still be interesting to look at in order to get an idea of how consistent the sensors were relative each other. Prior to analyzing the friction data, all surfaces that were misclassified were removed from the dataset.

The friction coefficients on dry (green markers), wet (red markers) and icy (blue markers) plates were plotted in Figure 2a. In most cases, the friction values seemed reasonable. The friction for a dry road was in the range 0.7-0.8, as water was added the friction was reduced somewhat (except for MetRoad Mobile), and as the water-film thickness was increased the friction was reduced to values between 0.4 and 0.7. On the icy plates, the MARWIS stood out by clearly giving too high friction coefficients, from 0.5 to 0.65. For the other sensors, the range of friction coefficients (0.1 to 0.3) are all within literature values for ice. However, looking at the breaking distances corresponding to the sensor friction coefficients

(Figure 2b), it is clear that these small differences in friction values at the low end of the scale actually corresponds to a huge range in the actual driving conditions. The calculated breaking distance (assuming a speed of 60 km/h) gives breaking distances on ice from 50 to 150 m. Having different sensors indicating such large differences in driving conditions on the same surface is clearly problematic, and it means that friction estimates on ice should be interpreted with caution. Considering that there are other, more reliable means to measure friction, these might be preferred. However interpreted with caution, especially in low friction conditions, optical road condition sensors can be a useful supplement to more accurate methods for measuring friction.

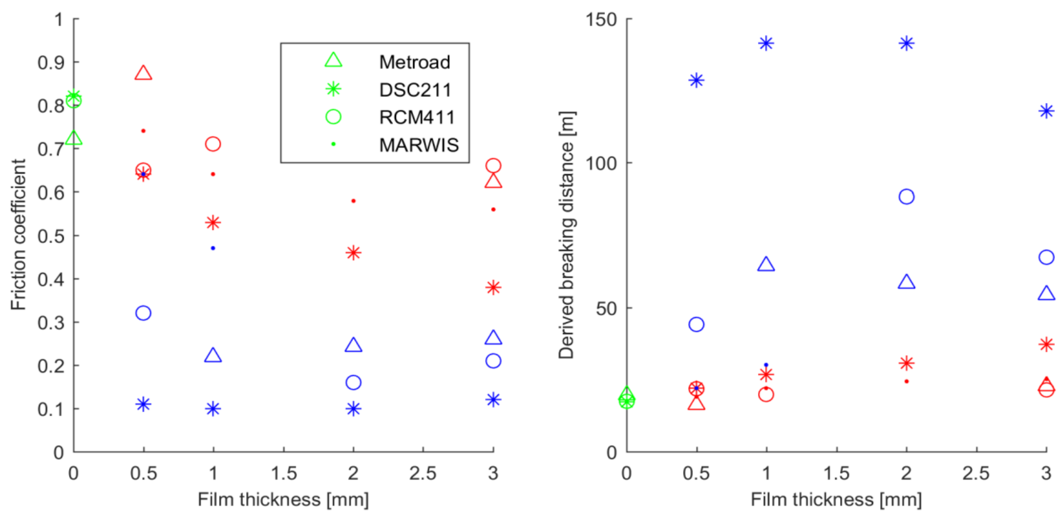

Figure 2 a) Friction coefficient provided by the different sensors on films of water (red markers) and ice (blue markers). The dry asphalt reference is given as a 0 mm film thickness (green markers). b) Shows the derived breaking distance from the acquired friction coefficients, assuming a velocity of 60 km/h.

The friction values on snow ranged between 0.25 and 0.4 (Figure 3), values well within the limits found in literature, with MARWIS and RCM411 giving somewhat lower values than the other sensors. The sensors, with the exception of RCM411, gave more or less constant friction values irrespective of snow type. As was pointed out in the classification-results section, this is clearly not correct as the very compact fresh snow is expected to have a much lower friction than the uncompacted fresh snow. The RCM411 did show a slightly lower friction value for the highly compressed snow than for the uncompressed snow. However, the RCM411 friction on old compacted snow and on spring-snow was even lower, a results which seems doubtful and the result should be validated in the field. It seems like all sensors treat snow as being one material, with one set of properties. Considering the large range of snow types and driving conditions they can provide, this is a limitation to the use of optical sensors for friction estimates.

On snow with added gravel (Figure 4), only the MARWIS showed some kind of improvement in friction as more gravel was added, from 0.19 on pure snow to 0.35 for snow with 25% gravel. The RCM411, on the other hand, showed a decrease in friction with an increased gravel amount. No other sensor derived friction was significantly affected by the amount of gravel. The reason was probably identified in the classification-results section; that there is no specific class for gravel. Since the sensors does not recognize gravel, they also do not take this into account when calculating a friction value. The reason the RCM411 showed a decreased friction was likely because it classified the gravel as ice.

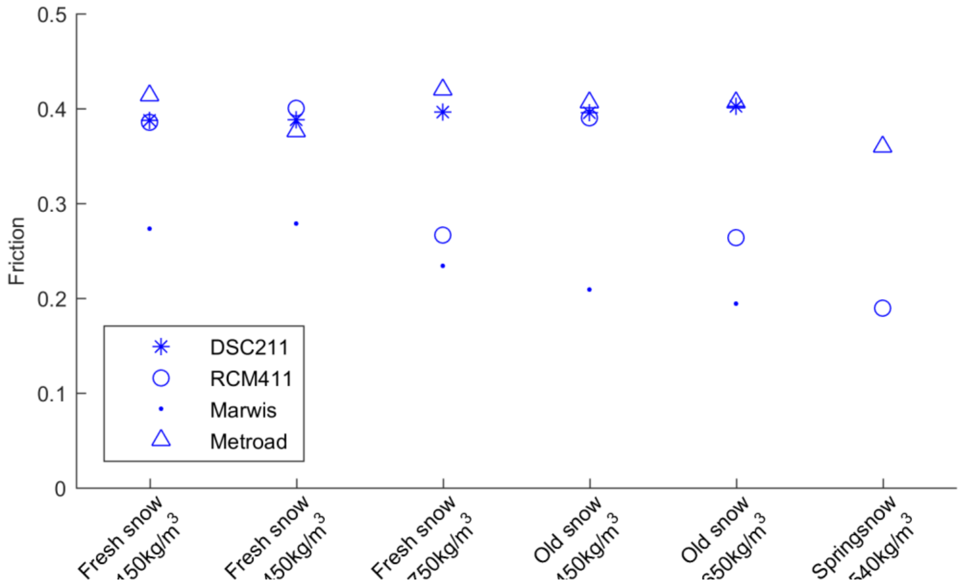

Figure 3 The derived friction values on the different types of snow.

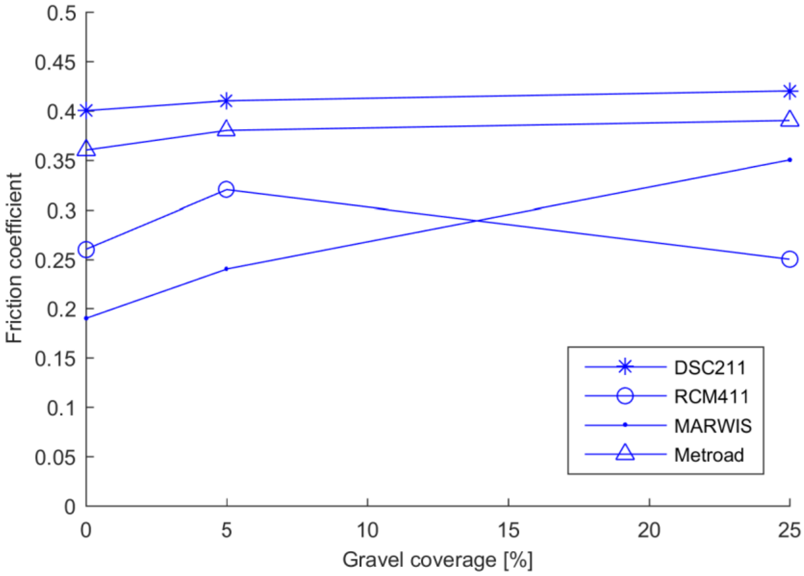

Figure 4 The change of derived friction coefficient as a function of the amount of gravel coverage on compacted snow.

3.3 Water film thickness

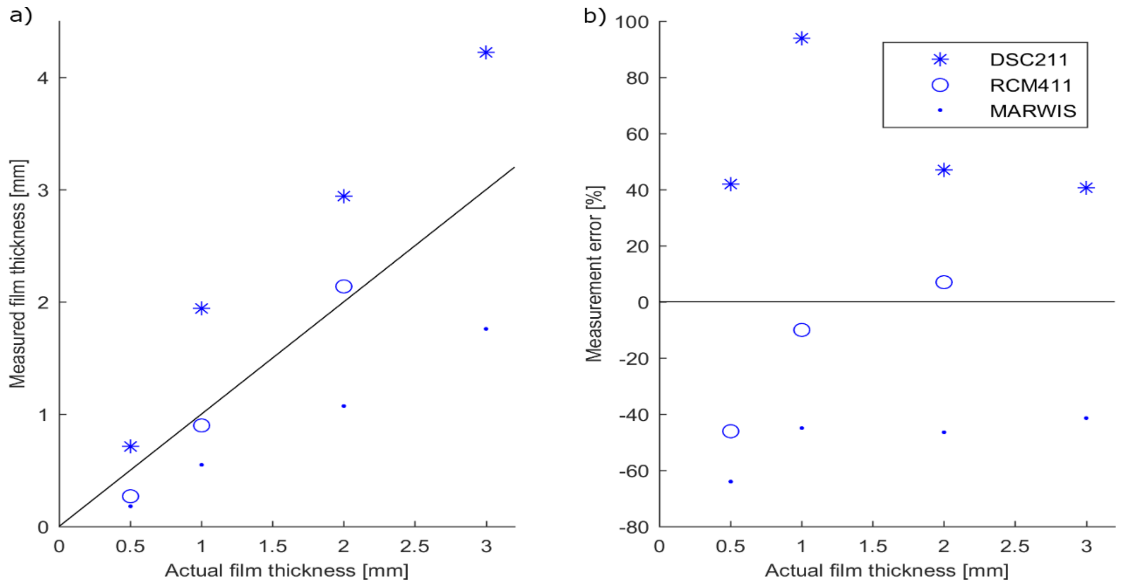

Figure 5. The water film thickness on the gray asphalt measured by three of the sensor, a) as a function of actual water film thickness and b) the corresponding error of the measurements.

Table 3 Comparison of the measured water film thickness of the mobile sensors on the gray and black asphalt

Sensor	/ Actual film thickness	1 mm	2 mm	3 mm
MARWIS	Gray	0.55 mm	1.07 mm	1.76 mm
	Black	0.27 mm	0.59 mm	0.9 mm
Teconer	Gray	0.9 mm	2.14 mm	Saturated
	Black	1.79 mm	Saturated	Saturated

Water film thickness is only measured by three of the tested sensors; the DSC211, the RCM411 and the MARWIS. The results (Figure 5) showed that all sensors could register increased amounts of water on the substrate. The absolute thickness, however, was highly variable between the sensors with results varying with a factor of 2 - 3. Naturally, this meant that the accuracy was limited, correspondingly. MARWIS consistently measured too shallow, with an error of -40 to -60 %. The DSC211 constantly measured too high, with errors in the range of 40-100%. The RCM411 measured too shallow for the thinner films (>2mm), but too high for the thicker films. The RCM411 data point at 3 mm is not included, as the sensor saturated before reaching this thickness. Replacing the gray asphalt with the black had a dramatic effect on the measured water film thickness (Table 3). For MARWIS the measured value decreased to half of that on the gray plate, increasing the error from minus 40%-60% to minus 80-100%. For the RCM411, on the other hand, the measured film thickness was doubled, increasing the error from 10-40% to near 100%. Overall, water-film thickness seems to be much less reliable than the sensor resolution implies, and users should be aware of this when interpreting the results. However, considering the large span of film-thicknesses on roads (30 µm – 3 mm) and the fact that there are few alternatives when it comes to measuring the water-film thickness on roads, an accuracy of a factor two to three may be adequate.

3.4 Ice thickness and snow-water equivalence

Only the Vaisala DSC211 attempts to measure the ice film thickness and the snow-water equivalence. The ice film thickness measurements (Figure 6) shows that the DSC211 performs well in estimating ice film thickness in the range from 0.5 to 3.5 mm. Measurement errors are less than 25% in the tested thickness range. This is something which could be very useful in a deicing situation, where knowing the ice layer thickness and the temperature makes in possible to calculate the amount of salt required to melt the ice.

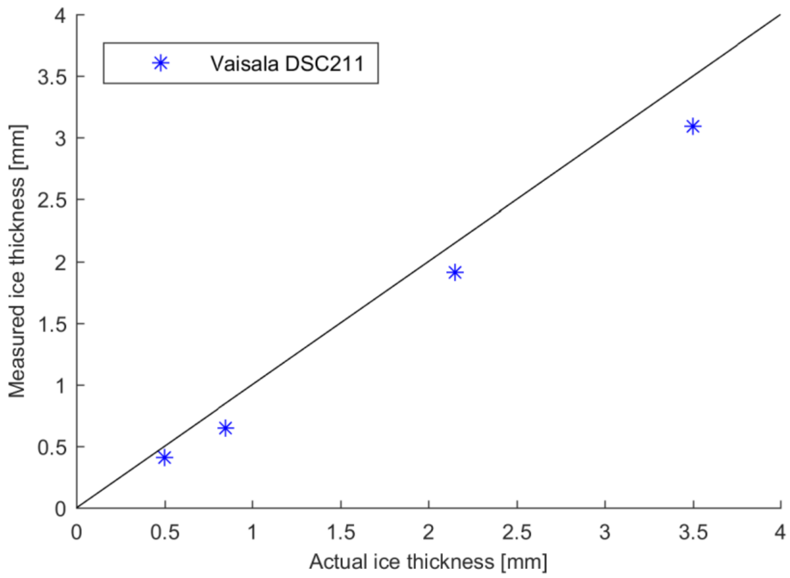

Figure 6 Ice film thickness measured by the DSC211. The black line indicates a 0% measurement error.

The snow-water-equivalence (Figure 7) was spot-on for the highly compacted fresh snow. In general, the algorithm performed well for the fresh snow, with errors within ± 100%. For the older snow types, irrespectively of the density, the DSC211 snow-water-equivalence measurements were significantly worse with errors of 200% or more. Having an algorithm which is tuned for compacted fresh snow makes very much sense for salted roads, where the snow is removed before it has time to age. On non-salted winter roads, however, this means that when interpreting the data it is good to know that the snow cover thickness likely will be underestimated by as much as a factor 3.

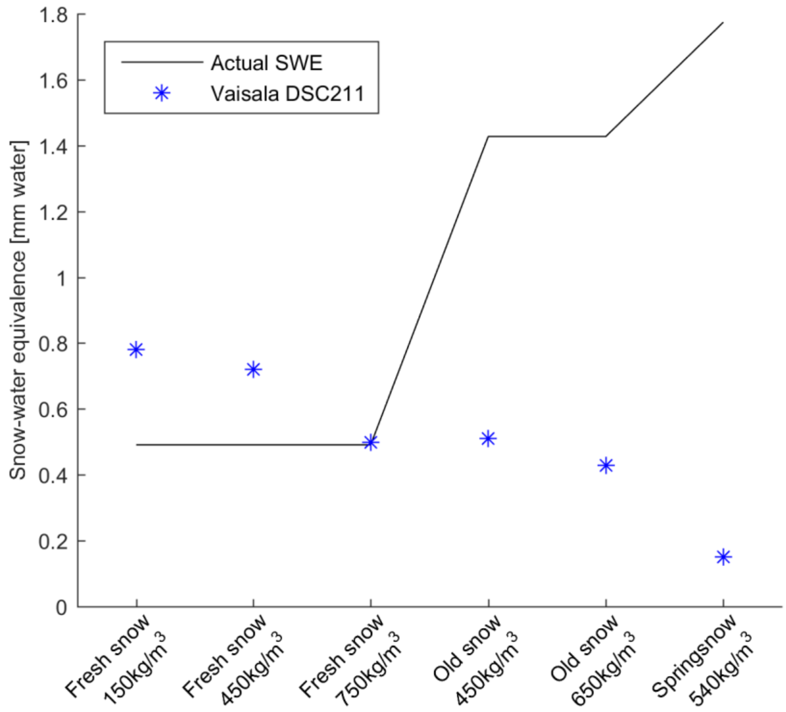

Figure 7 Measured snow-water-equivalence for the different snowtypes and densities

3.5 Ice percentage

The only sensor giving ice percentage, which is the amount of ice on the road surface, as an output was the Lufft MARWIS. This algorithm was only really relevant in the gravel-amount test. The ice-percentage algorithm of the MARWIS responded to an increased amount of gravel, and as the actual gravel percentage increased from 0% to 5% and 25%, the MARWIS ice percentage showed 96% and 80% ice respectively. This shows that the ice percentage algorithm of the MARWIS works well. With an additional gravel-classification, this could become very useful on non-salted winter roads.

5. Conclusions

I'll leave this for ROSTMOS, since I believe you are better suited for the task.

6. Acknowledgements

Thanks to everyone involved. I leave this too, since I believe you know better than me whom to thank.

7. References

Brown, B. & Baass, K., 1997. Seasonal Variation in Frequencies and Rates of Highway Accidents as Function of Severity. *Transportation Research Record: Journal of the Transportation Research Board*, 1581(970491).

Casselgren, J., Sjödahl, M. & Leblanc, J., 2007. Angular spectral response from covered asphalt. , 46(20), pp.4277–4288.

Jonsson, P., Casselgren, J. & Thörnberg, B., 2015. Road Surface Status Classification Using Spectral Analysis of NIR Camera Images. , 15(3), pp.1641–1656.

Kinosita, S. et al., 1970. Classification of snow and ice on roads. *Highway Research Board Special*

Report, 115(115), pp.8–16. Available at: http://pubsindex.trb.org/view.aspx?id=110808 [Accessed August 8, 2011].

Kuehnle, A. & Burghout, W., 1998. Winter Road Condition Recognition. *Transportation Research Record: Journal of the Transportation Research Board*, 1627(98), pp.29–33.

Meindl, T. et al., 2006. An Embedded Hardware-Software System to Detect and Foresee Road Ice Formation. In *International Jount Conference on Neural Networks*. Vancouver, Canada, pp. 4884–4891.

Räma, P. & Kulmala, R., 2000. Effects of variable message signs for slippery road conditions on driving speed and headways. *Transportation Research Part F: Traffic Psychology and Behaviour*, 3, pp.85–94.

Rämä, P., 1999. Effects of Weather-Controlled Variable Speed Limits and Warning Signs on. *Transportation Research Record*, 1689(99), pp.53–59.

Schleef, S. et al., 2014. An improved machine to produce nature-identical snow in the laboratory. *Journal of Glaciology*, 60(219), pp.94–102. Available at: http://www.igsoc.org/journal/60/219/t13J118.html [Accessed January 9, 2014].

Usman, T., Fu, L. & Miranda-Moreno, L.F., 2010. Quantifying safety benefit of winter road maintenance: Accident frequency modeling. *Accident Analysis and Prevention*, 42(6), pp.1878–1887. Available at: http://dx.doi.org/10.1016/j.aap.2010.05.008.

Viikari, V. V., Varpula, T. & Kantanen, M., 2009. Road-condition recognition using 24-GHz automotive radar. *IEEE Transactions on Intelligent Transportation Systems*, 10(4), pp.639–648.