
MSURJ

MCGILL SCIENCE UNDERGRADUATE RESEARCH JOURNAL
VOLUME 14 \ ISSUE 1 \ APRIL 2019

ON THE COVER

The Eastern North Pacific is a region of climatologically significant cyclone activity, often associated with extreme weather in North America. On page 12, Yeechian et al. investigate the influences of middle-tropospheric weather patterns on downstream storm activity.

The cover image illustrates peak weather regime intensities over the eastern Northern Pacific.

BOARD 2018-2019

Editors-in-Chief

Sebastian Andric
Neuroscience

Howard Li
Biochemistry

Managing Editors

Abtin Ameri
Mechanical Engineering

Deifilia To
Chemical Engineering

Graydon Tope
Biology and Anthropology

Jacqueline Yao
Microbiology and Immunology

Senior Editors

Yingke Liang
Biochemistry

Richard Yang
Biochemistry

Shannon Egan
Physics

Janet Wilson
Physiology

Carrie Gu
Pharmacology

Mathilde Papillon
Physics

Editors

Laura Meng
Neuroscience

Janet Tang
Microbiology and Immunology

Viet Hoang
Biology and Computer Science

Katharine Kocik
Biology and English

Advisor

Victor Chisholm
Office for Undergraduate Research in Science

Layout Designers

Abtin Ameri

Carrie Gu

Laura Meng

Viet Hoang

Graphic Designer

Julia Luo
Microbiology & Immunology

805 Sherbrooke St. West, Room 1B22
Montreal, Quebec, H3A 2K6 Canada
Phone: (514) 398-6979
Fax: (514) 398-6766
Email: msurj.sus@mail.mcgill.ca
Website: msurj.com

TABLE OF CONTENTS

7 Foreword

9 Acknowledgements

RESEARCH ARTICLES

12 Winter Mid-Tropospheric Weather Regimes in the Eastern North Pacific
Yeechian Low et al.

18 Variability and Predictability of the Bering Strait Ocean Heat Transport and Arctic Ocean Sea Ice Extent
Jed Lenetsky et al.

24 Predator Driven Trait Changes in *Anolis smaragdinus*
Samantha Wunderlich

28 Examining the Transition from a perennial to a seasonal sea ice cover in the Arctic Ocean: A Lagrangian Approach
Jamie Hart et al.

34 Habitat mapping of giant kelp (*Macrocystis pyrifera*) and devil weed (*Sargassum horneri*) off the coast of Santa Catalina Island, California
Michael Espriella et al.

40 Physical Responses to Anthropogenic Disturbance on a Runway in the Canadian High Arctic
Tabatha Rahman et al.

REVIEW ARTICLES

46 The Bionic Human: A Review of Interface Modalities for Externally Powered Prosthetic Limbs
Nipun Mundkur

FOREWORD

Dear Reader,

Science, by definition, is an evolving discipline; and this is now more true than ever. With each passing moment, science expands further and faster than ever before, tirelessly extending the scope and depth of human knowledge. Meanwhile, the research community itself has also changed. Gone are the days where major breakthroughs are achieved by a select few. While the work of these revolutionary individuals built the foundations upon which we stand today, the complex problems facing us require the collective effort of teams of scientists collaborating globally with one another. It is in this context that undergraduate research has been able to flourish with students engaging in impactful work to tackle our toughest challenges.

When the McGill Science Undergraduate Research Journal was launched fourteen years ago, we were one of the first organizations in North America to celebrate and publish peer-reviewed undergraduate research. Today, reflecting the prevalence of undergraduate work, undergraduate research journals have become a staple in many academic institutions.

This year, we have received more high quality papers than ever before from not only McGill University but all over the world. The words and figures in these pages represent the incredible hard work and passion of all of our authors. We are extremely proud to be able to present these seven articles to you this issue and we hope that we are able to continue to do so for many years to come.

On behalf of the editorial board, thank you.

Howard Li & Sebastian Andric
Editors-in-Chief

ACKNOWLEDGEMENTS

The McGill Science Undergraduate Research Journal would like to thank its generous contributors without whom this journal would not have been possible.

We thank Dean Bruce Lennox who played a pivotal role in helping MSURJ set new goals for the upcoming decade, without his advice and expertise, MSURJ would not be the success it is currently. We would also like to thank the McGill Librarians for their help in transitioning MSURJ to an online electronic system. Lastly, we would like to acknowledge Mr. Victor Chisholm for his continuous support and guidance.

We thank all the donors in the McGill community for their generous support:

Faculty of Science
Faculty of Medicine
Department of Pharmacology
Department of Geography
Department of Chemistry
Science Undergraduate Society of McGill (SUS)
Chemistry Undergraduate Student Society (CUSS)

We extend our gratitude to the numerous peer reviewers who took the time to review all of the submissions. Lastly, we wish to recognize the student contributors whose efforts have made the journal possible.

Research Article

¹Department of Atmospheric and Oceanic Sciences, McGill University

Keywords

500-hPa geopotential height, Mid troposphere, Weather regime

Email Correspondence

yeechian.low@mail.mcgill.ca

Yeechian Low¹, Eyad Atallah¹, John Gyakum¹

Winter Mid-Tropospheric Weather Regimes in the Eastern North Pacific

Abstract

Background: The eastern North Pacific (ENPac) is a region of climatologically significant cyclone activity, often associated with extreme weather in North America. Regions of high (ridges) and low (troughs) 500-hPa height typically drive this activity. We identify 500-hPa height time variability extremes as “regimes.” Our objectives are to determine the regimes’ characteristics, predictability, and relationships to North American extreme weather.

Methods: We define weather regimes, separating them into two types based on whether the 500-hPa height variance is extremely low or high. We analyze their general characteristics during the winter (December, January, and February) and relationships to extreme North American weather. To analyze the regimes’ predictability, we define forecast discontinuities as significantly improved extreme 500-hPa height variability model forecasts compared with model forecasts verifying at the same time, but initialized 24 hours earlier. We analyze their effects on anticipated weather.

Results: ENPac low variance regimes are usually dominated by one or two large, slow-moving features, usually a trough with an associated surface cyclone 200-700 km to the west and a ridge with an associated surface anticyclone 200-700 km to the east. This pattern leads to anomalous southerly winds and moisture transport. Low variance regimes are generally associated with anomalous wetness in northwestern Canada, warmth in western North America, and dryness in the southwest U.S. High variance regimes are usually dominated by smaller, faster-moving features that alter the 500-hPa heights substantially. These regimes are more varied, but there is a tendency to have a ridge 200-700 km to the west and a trough 200-700 km to the east, leading to anomalous northerly winds and transport of drier polar air into the ENPac region. High variance regimes are generally associated with anomalously cold air in western North America and wetness in the western U.S. Some forecast discontinuities are associated with changes in anticipated weather locally in the ENPac region, while other discontinuities are associated with changes in anticipated weather on a much larger scale, extending to North America.

Limitations: Limitations include the small sample of regimes found during the period of record (18 low variance and 10 high variance), the metric being limited to the 500-hPa level, and the study of only the ENPac winter.

Conclusion: Low and high variance regimes generally lead to different ENPac weather patterns and North American extreme weather. Forecast discontinuities differ significantly from each other in their spatial extents. Further work is necessary to identify their causes and characteristics.

Introduction

The eastern North Pacific is a region of climatologically significant cyclone activity (1), often associated with extreme weather downstream in North America, such as an extreme cold-air outbreak. Middle-tropospheric weather patterns are key drivers of this storm activity (1). By characterizing and studying them in further detail, we may be able to better predict these patterns and their impacts further in advance. A better understanding of why certain model forecasts simulate certain weather patterns better than others may help give a better sense of predictability.

Many previous studies (e.g., (2), (3), and (4)) focused on weather patterns or regimes characterized by persistent, quasi-stationary features such as atmospheric blocks—a nearly stationary, persistent ridge that redirects or splits atmospheric flow in the mid-upper troposphere around it. Locations underneath a block usually experience persistent dryness, while areas on the periphery of the block experience persistent storminess. However, there is no widely used single definition of a block or weather regime. In (2), blocking is defined based on the presence of certain upper-level flow characteristics; upper-level charts would be inspected by a researcher to determine if there was blocking. The areas of most frequent occurrence, characteristic movement and persistence, and seasonal and yearly trends of blocking activity are determined. In (3), a zonal index suitable for identification of blockings is defined and translated into a computer program, and the characteristics of Northern Hemisphere blocking situations are assessed. In (4), a slightly modified version of the objective zonal index

used in (3) is used to quantify both observed and forecasted occurrence of blocking. To expand the scope of studies of weather regimes, we produce a new objective definition of regimes, split into two categories: Low variance regimes, where the synoptic-scale weather pattern changes relatively little, compared to climatology; and High variance regimes, where the synoptic-scale weather pattern changes substantially, compared to climatology.

We also analyze weather patterns associated with extreme North American weather during the regimes and objectively define and analyze forecast discontinuities, which are significantly improved forecasts of 500-hPa geopotential height—the height where the atmospheric pressure is 500 hPa, which is about half that of sea level—variability during regimes compared with forecasts verifying at the same time, but initialized 24 hours earlier. We analyze the relationship between these forecast discontinuities and sensible weather, both locally in the ENPac region and more broadly over North America.

Definitions

500-hPa geopotential height: the height where the atmospheric pressure is 500 hPa, approximately half that of sea level.

Mid-troposphere: around 5-6 km altitude, often indicated by the altitude of the 500-hPa level

Weather regime: a large-scale recurrent atmospheric flow pattern with various specific definitions depending on study

Methods

Defining a Regime

We define regimes based on the variability of the 500-hPa height, a mid-tropospheric level that is typically at around 5 km above sea level. We use the National Centers for Environmental Prediction (NCEP) reanalysis 2 dataset ($2.5^\circ \times 2.5^\circ$ resolution) (5) to calculate the areally-averaged 500-hPa height over the ENPac region during the winters (December, January, and February) of 1979–1980 to 2015–2016. As shown in Fig. 1, we choose the ENPac region such that it is in the left exit region of a climatological North Pacific jet stream, which is a favorable region for surface cyclogenesis, or the formation or strengthening of a surface low-pressure system. A jet stream is an elongated zone of very strong westerly winds at the upper-levels of the atmosphere.

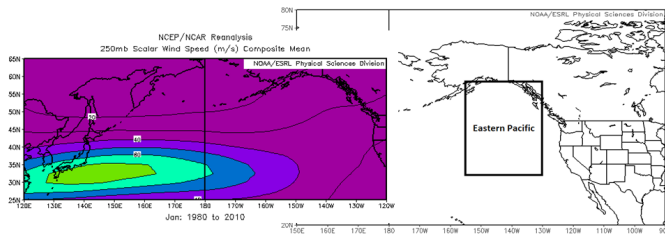


Figure 1. Left: Composite mean of 250-hPa wind speed (m/s) in January from 1980–2010. Right: The ENPac region we chose.

In this case, the synoptic-scale variability, rather than the seasonal variability, of the 500-hPa height is of interest, so we perform linear regression on the seasonal 500-hPa height and then subtract the resulting line from the 500-hPa height to obtain the de-trended 500-hPa height for each season. We then compute the 7-day running standard deviation of the de-trended areally-averaged 500-hPa height and its daily climatology. We obtain the standard deviation anomaly by subtracting the daily climatology from the standard deviation. For brevity, we call this final quantity the *height variance*. We use this metric to capture the departure of the 500-hPa variability from the climatological average. We define a *regime* as a period of at least five days during which the height variance is continuously below the 10th percentile (low variance regime) or above the 90th percentile (high variance regime). Fig. 2 shows examples of such regimes.

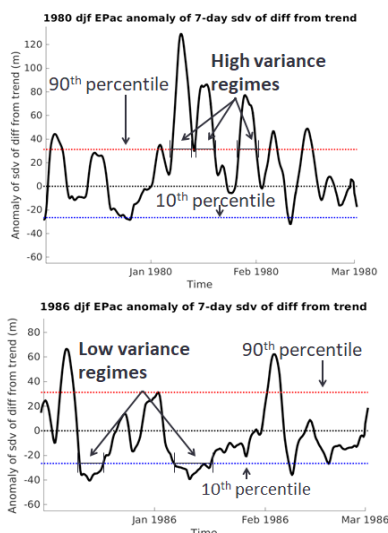


Figure 2. Height variance (anomaly of standard deviation in meters) in winter 1979–1980 (top) and winter 1985–1986 (bottom) in the ENPac region, with the regimes marked.

Analyzing Weather Patterns During Regimes

To analyze the average weather pattern in the ENPac region during low and high variance regimes, we calculate composites (averages) of various metrics during the times at which low and high variance regimes are

at peak intensity. These metrics include anomalies of 500-hPa height, 250-hPa meridional wind (the 250-hPa level is typically at around 11 km altitude), precipitation rate, and mean sea-level pressure (MSLP), as well as the composite MSLP, thickness, and thickness anomaly map. The peak of a (low/high) regime occurs when the height variance is the (lowest/highest). The map consists of MSLP in solid black lines for every 8 hPa, 1000–500 hPa thickness in green dashed lines for every 60 m, and 1000–500 hPa thickness anomaly in meters shaded. We also construct plots on individual days during regimes to show the time-evolution of regimes and what patterns can be found at any given time during a regime (not shown).

To analyze extreme weather over North America, we use the North American Regional Reanalysis (NARR) dataset (6), a much higher-resolution dataset ($0.3^\circ \times 0.3^\circ$ resolution) specifically for North America and surrounding areas. In this study, we define North America to be the land area from 20–72°N and 50–165°W. From the NARR, we derive three variables to identify extreme weather: daily average 2-meter temperature, daily average 850-hPa equivalent potential temperature (θ_e), and daily precipitation. The 850-hPa θ_e is an atmospheric variable describing how warm and humid the air is at the height where the atmospheric pressure is 850 hPa, at approximately 1.5 km altitude. θ_e is a conserved variable for air parcels in the absence of heat exchange with the environment, and 850-hPa corresponds approximately to the base of the precipitation layer. We use the 850-hPa θ_e in addition to the 2-meter temperature to assess how warm and moist the air through a deeper layer of the atmosphere, not just near the surface. The formula for θ_e is given in the following equation (Eq. 1):

$$\theta_e \approx \left(T + \frac{L_v}{c_{pd}} r \right) \left(\frac{p_0}{p} \right)^{\frac{R_d}{c_{pd}}}$$

where T is temperature, L_v is the latent heat of evaporation, c_{pd} is the specific heat of dry air at constant pressure, $p_0 = 1000$ hPa is the standard reference pressure, p is the pressure at the point, r is the mixing ratio of water vapor, and R_d is the specific gas constant for dry air. The extreme limits are 10th and 90th percentiles for 2-meter temperature and 850-hPa θ_e , and 90th percentile of daily precipitation ≥ 0.2 mm. We calculate the percentiles for each grid point based upon the monthly climatology (e.g. one value for the 10th percentile in January, another value for February, etc) using the NARR data from the winters of 1979–1980 to 2015–2016. Then, we compute the geographic areas and areal percentage of North America for extremes of each of the three variables on a daily basis for all winters from 1979–1980 to 2015–2016.

Defining and Analyzing Forecast Discontinuities

To assess the predictability of the ENPac regimes, we also study instances in which a model run is appreciably better (closer to actual) than the previous model run that was initialized 24-h earlier, using the Global Ensemble Forecast System (GEFS) reforecast model ensemble ($1.0^\circ \times 1.0^\circ$ resolution) (7). The model ensemble consists of one “original” control run and ten ensemble members. Ensemble members are simulations run, each with a slight variation of its initial conditions from the control run, to convey the forecast uncertainty caused by imperfect initial conditions. The reforecast is run once a day initialized at 0000 UTC. So far in this study, we focus on the control run. We define a *forecast discontinuity* as an instance where the GEFS reforecast control run height variance is at least 30 m for a low variance regime (or 45 m for a high variance regime) closer to the actual height variance than the previous run averaged over a 2-day period. These values correspond roughly to the 90th percentile of height variance changes modeled by the GEFS control run. We use a higher threshold for high variance regimes because modeled height variances are more variable for high variance regimes than low variance regimes. We compare the previous run, subsequent run, and reanalysis that are valid at the same time for the discontinuities.

Results

ENPac Weather Patterns During Regimes

Using the criteria specified in the previous section, we find 18 low vari

ance regimes and 10 high variance regimes (Tab. 1). In low variance regimes, there are on average anomalously low 500-hPa heights, typically associated with a surface cyclone 200-700 km west of the ENPac region, and anomalously high 500-hPa heights, typically associated with a surface anticyclone 200-700 km east of the ENPac region (shown in Fig. 3a and Fig. 3g). This pattern leads to anomalous southerly winds and moisture transport in the ENPac region (shown in Fig. 3d and Fig. 3f). The structure and exact placement of these features depends on the regime: on one extreme, the 14 Dec-21 Dec 2010 regime had the surface cyclone right in the ENPac region; while at the other extreme, the 10 Feb-17 Feb 2013 regime had it around 1000 km to the west. The average for high variance regimes is roughly the opposite, with anomalously high 500-hPa heights 200-700 km west of the ENPac region and an associated surface anticyclone in the ENPac region, as well as a trough 200-700 km east of the ENPac region (shown in Fig. 3b and Fig. 3h). However, for high variance regimes, the patterns are more varied, and this “average” setup usually is only present for a part of a regime due to the rapid changes in the weather pattern. For example, in the 7 Jan-13 Jan 1980 high variance regime, a strong surface anticyclone to the north of the ENPac region moved westward and cold air moved into the ENPac region from northwestern Canada, while in the 12 Jan-17 Jan 2008 high variance regime, there was a quick succession of surface cyclones and anticyclones moving eastward through the ENPac region.

Regime Classification	Regime Cases	Notable Weather (if any)
Low	30 Nov 1984 - 8 Dec 1984	Rossby wave break and buildup of anomalously high available potential energy period (11)
Low	6 Jan 1985 - 12 Jan 1985	Pineapple express 3 days after (9)
Low	11 Dec 1985 - 17 Dec 1985	Pineapple express 3 days after (9)
Low	6 Jan 1986 - 16 Jan 1986	Northwest U.S. floods 1 day after (10)
Low	5 Feb 1988 - 14 Feb 1988	
Low	22 Dec 1989 - 27 Dec 1989	Cold air generation 1-5 days before (12)
Low	28 Dec 1998 - 2 Jan 1995	
Low	12 Jan 1996 - 18 Jan 1996	Pineapple express (8)
Low	9 Dec 1997 - 14 Dec 1997	Rossby wave break and buildup anomalous period 4-14 days after (11) and pineapple express (9)
Low	5 Feb 1998 - 14 Feb 1998	
Low	24 Nov 2002 - 5 Dec 2002	Pineapple express (9)
Low	20 Jan 2005 - 26 Jan 2005	
Low	9 Feb 2007 - 16 Feb 2007	
Low	14 Dec 2010 - 21 Dec 2010	
Low	12 Dec 2011 - 18 Dec 2011	
Low	14 Feb 2012 - 19 Feb 2012	
Low	10 Feb 2013 - 17 Feb 2013	
High	7 Jan 1980 - 13 Jan 1980	Deep cold air generation over northwestern Canada (13)
High	14 Jan 1980 - 19 Jan 1980	
High	27 Jan 1980 - 1 Feb 1980	Deep cold air generation over northwestern Canada (13)
High	7 Feb 1982 - 13 Feb 1982	
High	23 Dec 1982 - 27 Dec 1982	Long duration freezing rain (C. McCray, 2018, personal communication) and extreme θ_e for Montreal (14)
High	28 Jan 1991 - 2 Feb 1991	
High	16 Jan 1993 - 23 Jan 1993	
High	29 Nov 2007 - 7 Dec 2007	
High	12 Jan 2008 - 17 Jan 2008	Extremely high 850-hPa θ_e for Montreal 3 days before (14)
High	8 Feb 2011 - 13 Feb 2011	Rossby wave break and buildup of anomalously high available potential energy (11)

Table 1. All regime cases with notable weather found in previous research if any.

At any given time during a low variance regime, there tends to be one or two large, slow-moving features that make the 500-hPa height field relatively steady in the ENPac region with small fluctuations. On the other

hand, at any given time during a high variance regime, there tends to be faster-moving features that change the 500-hPa height field substantially in the ENPac region.

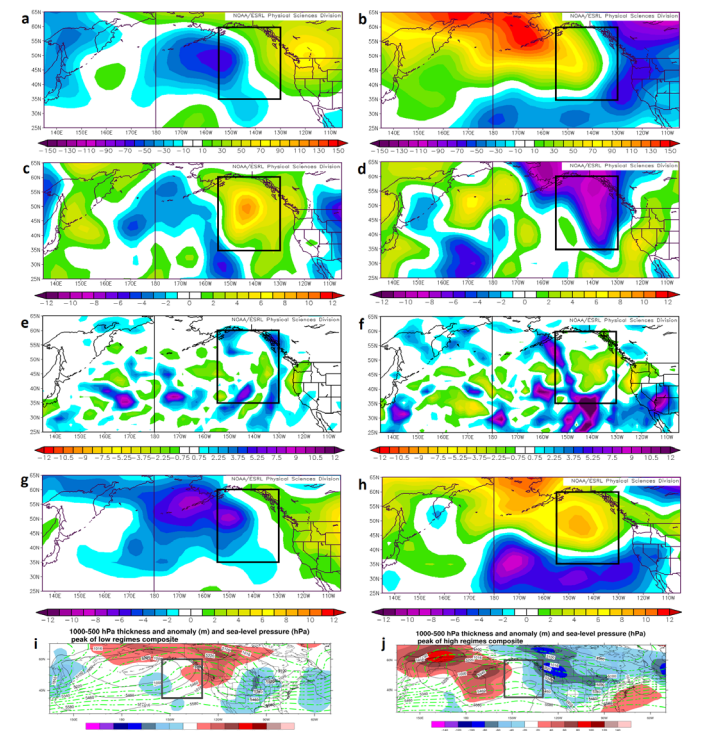


Figure 3. Anomalies of various meteorological fields composited at the time of peak regime intensity during low variance regimes (left) and high variance regimes (right) with the ENPac region boxed. a)-b): anomalies of 500-hPa height (m); c)-d): 500-hPa meridional wind (m s⁻¹); e)-f): precipitation rate (mm day⁻¹); g)-h): mean sea level pressure (hPa); i)-j): MSLP, thickness, and thickness anomaly map.

North American Extreme Weather During Regimes

As shown in Fig. 3i (low variance regime composite), the counterclockwise circulation around the surface cyclone in the Central Pacific brings southerly winds and anomalous warmth measured by anomalously high 1000-500 hPa thickness near the West Coast of North America. (The 1000-500 hPa thickness is a measure of the average temperature of the air column between the surface and 500 hPa.) The southerly winds sometimes bring moisture plumes northward into the West Coast and southern Alaska, which in conjunction with the surface cyclone in the central Pacific, causes heavy precipitation. Meanwhile, the southwest U.S. is drier as the upper-level ridge near the West Coast (shown in Fig. 3a) deflects most of the moisture and surface cyclones to the north and west. One specific case (a day in the 6 Jan-16 Jan 1986 regime) is shown in Fig. 4a-4d, which shows deep warm air (shown in very high 2-meter temperature, 850-hPa theta-e, and 1000-500 hPa thickness) over parts of western Canada and heavy precipitation in British Columbia. Different regimes have somewhat different regions affected by the moisture plumes; for example, in the 6 Jan-12 Jan 1985 regime, the precipitation was directed from southern Alaska to northern British Columbia, while in the 6 Jan-16 Jan 1986 regime, the precipitation was directed into British Columbia and Washington.

As shown in Fig. 3j (high variance regime composite), the clockwise circulation around the weak surface anticyclone near the middle of the ENPac region brings northerly winds and anomalous cold measured by anomalously low 1000-500 hPa thickness in western North America, especially in western Canada. The cold anomalies to the north and less cold anomalies to the south indicate that the meridional temperature gradient is anomalously strong. The upper-level trough in western North America (shown in Fig. 3b) allows surface cyclones to penetrate farther south and into the western U.S., sometimes leading to extreme precipitation there. Also, the southwest U.S. is relatively dry on average, so a heavy precipita-

tion event there is more extreme for that area than it would be for other parts of North America. A specific case (a day in the 7 Jan-13 Jan 1980 regime) is shown in Fig. 4e-4h, which shows deep cold air (shown in very low 2-meter temperature, 850-hPa theta-e, and 1000-500 hPa thickness) over a significant part of western North America and heavy precipitation in the western U.S. One particular exception is the 12 Jan-17 Jan 2008 regime mentioned earlier, which had warmer than normal temperatures in central and northern Canada and colder than normal temperatures in the southern U.S.

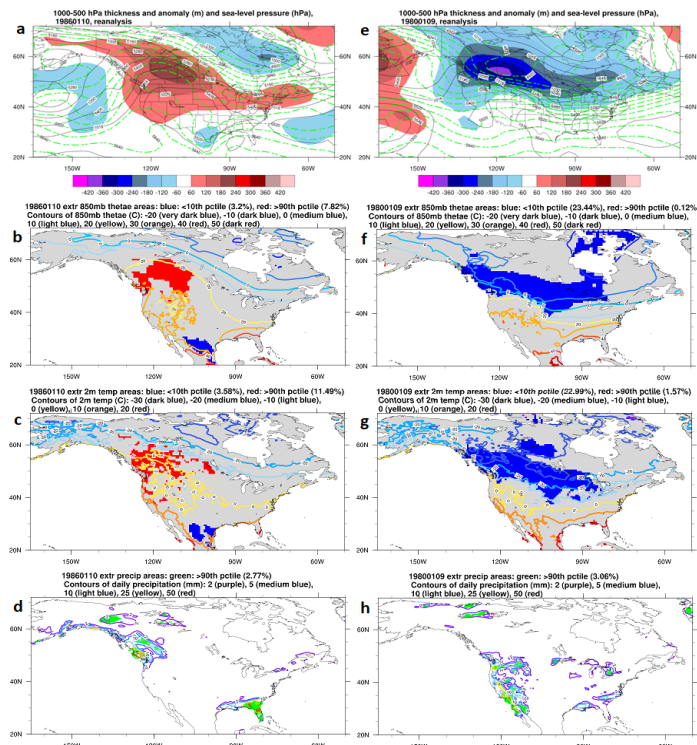


Figure 4. a)-d): Daily extreme weather metrics for 10 Jan 1986 low regime peak. a): MSLP, thickness, and thickness anomaly map. b): 2-meter temperature (contoured every 10 °C), extremes (blue–low, red–high). c): same as b) for 850-hPa theta-e. d): precipitation (contoured 2–100 mm), extremes (green). e)-h): same as a)-d) for 9 Jan 1980 high regime peak.

Forecast Discontinuities

Using the criteria specified in the previous section, we find 30 forecast discontinuities: 24 associated with low variance regimes and 6 associated with high variance regimes. Forecast discontinuities sometimes are associated with changes in anticipated weather over a large part of North America. For example, as shown in Fig. 5a-5c, the large-scale weather pattern over North America at 1200 UTC 31 Jan 1991 from the GEFS model forecast initialized on 24 Jan is almost the opposite of that shown by the forecast initialized on 23 Jan. The later run output is much closer to the reanalysis, or what actually happened. Other forecast discontinuities only have an impact on the anticipated weather conditions within the ENPac region. For example, as shown in Fig. 5d-5f, the large-scale flow pattern over North America at 1200 UTC 22 Jan 2005 from the GEFS model forecasts initialized on 14 and 15 Jan are very similar to each other. The main difference is the structure and position of the surface cyclone in the ENPac region, which was better predicted by the later run.

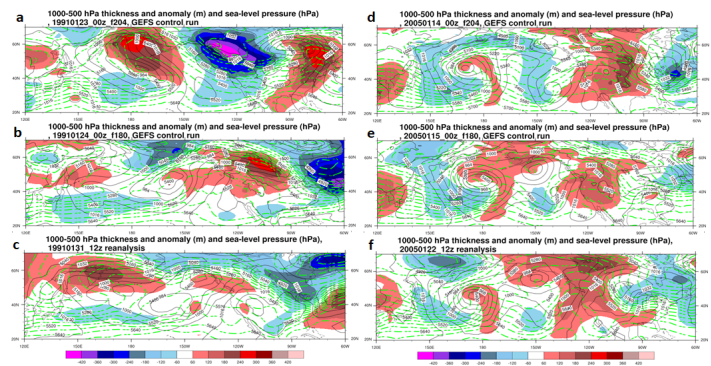


Figure 5. a)-c): MSLP, thickness, and thickness anomaly map of GEFS 23 Jan 1991 control run (a), GEFS 24 Jan 1991 control run (b), and reanalysis (c), all valid on 31 Jan 1991 at 1200 UTC. d)-f): same as a)-c) but for GEFS 14-15 Jan 2005 runs and valid on 22 Jan 2005 at 1200 UTC.

Discussion & Conclusion

Not surprisingly, low and high variance regimes can lead to different ENPac weather patterns and different types of high-impact weather over North America. Five of the low variance regimes (Tab. 1) are associated with a “pineapple express”, or a narrow, moisture-rich plume, extending into parts of western North America, leading to heavy precipitation events in those areas (8), (9). A sixth event is associated with northwest U.S. floods (10). This is consistent with a persistent southerly, moisture-rich flow in the ENPac and West Coast regions during low variance regimes.

In contrast, two of the high variance regimes are associated with deep cold air generation over northwestern Canada. The combination of this deep cold air generation and a generally northerly flow of polar air into the ENPac region is consistent with the tendency for cold air outbreaks to occur in western North America and the ENPac region during high variance regimes. Also, one of the high variance regimes is associated with a long duration freezing rain event, and two of them are associated with extremely high 850-hPa theta-e for Montreal (and its vicinity in southeastern Canada and the northeast U.S.), which is not found during any of the low variance regimes.

Freezing rain results from warm, moist air from the south rising and moving over cold, sub-freezing air to the north, which is easier with a stronger meridional temperature gradient. Extremely high 850-hPa theta-e in southeastern Canada and the northeast U.S. occurs during a normally brief surge of warm and moist air from the Gulf of Mexico, southeast U.S., and/or western Atlantic. Consequently, having anomalously warm temperatures in those regions increases the probability for the development of extremely positive theta-e anomalies in Montreal. On average, as shown in Fig. 3i and Fig. 3j, the Gulf of Mexico to western Atlantic region is warmer than normal during high variance regimes. Additionally, it is colder than normal in most of Canada, indicating an anomalously strong north-south temperature gradient. These factors are consistent with high variance regimes featuring an increased likelihood of extremely high 850-hPa theta-e in southeastern Canada and the northeast U.S., yet still maintaining enough cold air at low-levels for the potential for a long duration freezing rain event.

To date, we have not investigated in detail the causes of the forecast discontinuities, but we hypothesize that differences in cyclogenesis between the different model forecasts play a major role in producing forecast discontinuities. Forecast variability of cyclogenesis is high due to the complicated dynamics of cyclones and the higher sensitivity of cyclogenesis to the initial state of the atmosphere (15).

There are some noteworthy limitations of this study. The relatively small sample size of regimes (18 low variance and 10 high variance) and the relatively large variability between regimes in the same category precludes us from drawing too detailed or specific conclusions. Nonetheless, we believe the general statements in the results section are valid, since they are physically consistent, involve areas closest to the ENPac region, and have a

plausible meteorological explanation. Also, the fact that height variance at 500-hPa does not always directly impact the surface weather raises some caveats. Two model runs can have the same height variance yet have significantly different weather patterns. Two model forecasts can have very different height variances that could result from differences locally in the ENPac region and have similar larger-scale features over North America. Additionally, model forecasts can perform poorly in one area and well in another area. Finally, the study focuses on the ENPac winter; the characteristics of regimes and forecast discontinuities could be different for other seasons and regions.

Acknowledgements

This work was funded by two Science Undergraduate Research Awards (SURA), a NOAA NGGPS (#NA16NWS4680018) Grant, and a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant. Thanks to the McGill Faculty of Science and Dr. Rubin Gruber for funding the SURAs. Thanks to supervisor Prof. John Gyakum and to Dr. Eyad Atallah for guiding my project.

References

1. Gulev SK, Zolina O, Grigoriev S. Extra-tropical cyclone variability in the Northern Hemisphere winter from NCEP/NCAR reanalysis data. *Clim. Dynam.* 2001 July;17:795–809.
2. Rex, DF. Blocking Action in the Middle Troposphere and its Effect upon Regional Climate. *Tellus*. 1950 Nov;2-3:196-211.
3. Lejenäs H, Økland H. Characteristics of Northern Hemisphere blocking as determined from a long time series of observational data. *Tellus*. 1983 Oct;35A:350-62.
4. Tibaldi S, and Molteni F. On the operational predictability of blocking. *Tellus*. 1990 May;42A:343-65.
5. Kanamitsu M, Ebisuzaki W, Woollen J, Yang S-K, Hnilo JJ, et al. NCEP-DOE AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc.* 2002 Nov 1;83(11):1631–43. Dataset is available from: <https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis2.html>
6. Mesinger F, DiMego G, Kalnay E, Mitchell K, Shafran PC, et al. North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.* 2006 Mar 1;87:343–60. Dataset is available from: <https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html>
7. Hamill TM, Bates GT, Whitaker JS, Murray DR, Fiorino M, et al. NOAA's second-generation global medium-range ensemble reforecast dataset. *Bull. Amer. Meteor. Soc.* 2013 Oct 1;94:1553–65
8. Lackmann GM, Gyakum JR, Benoit R. Moisture transport diagnosis of a wintertime precipitation event in the Mackenzie River basin. *Mon. Wea. Rev.* 1998 Mar 1;126:668–91
9. Roberge A, Gyakum JR, Atallah EH. Analysis of intense poleward water vapor transports into high latitudes of western North America. *Wea. Forecasting*. 2009 Dec 1;24:1732–47
10. Lackmann GM, Gyakum JR. Heavy cold-season precipitation in the northwestern United States: Synoptic climatology and an analysis of the flood of 17–18 January 1986. *Wea. Forecasting*. 1999 Oct 1;14:687–700
11. Bowley K. Interpreting rapid increases of zonal available potential energy from a synoptic-scale perspective. 2017;Ph. D. Thesis, McGill University. 2017 Oct;158 pp.
12. Turner, JK, and Gyakum JR. The Development of Arctic Air Masses in Northwest Canada and Their Behavior in a Warming Climate. *J. Climate*. 2011 Sep 1;24:4618-33.
13. Bliankinshtein N. Radiative contribution to the formation and development of Arctic continental air masses over northwest Canada. M. Sc. Thesis, McGill University. 2018 Jun;87 pp.
14. Wood A. A Climatology of North American Air Masses and their Extremes in Montreal, Quebec. M. Sc. Thesis, McGill University. 2015 Jun;87 pp.
15. Mullen SL, Baumhefner DP. Monte Carlo simulations of explosive cyclogenesis. *Mon. Wea. Rev.* 1994 Jul 1;122:1548–67.

Research Article

¹McGill School of Environment,
McGill University²Department of Atmospheric and
Oceanic Science, McGill University

Keywords

Sea ice extent, Bering Strait, Ocean
Heat transport, Forecasting, Arctic

Email Correspondence

jed.lenetsky@mail.mcgill.ca

Jed Lenetsky¹, Bruno Tremblay², Charles Brunette²

Variability and Predictability of the Bering Strait Ocean Heat Transport and Arctic Ocean Sea Ice Extent

Abstract

Background: This study examines the monthly, seasonal, and interannual variations in Pacific Ocean heat transport entering the Arctic Ocean through the Bering Strait, and its influence on sea ice extent in the Arctic Ocean.

Methods: Monthly ocean heat transport is calculated using temperature and volumetric transport data from moorings deployed in the Bering Strait. Pearson correlations are calculated between the observed detrended monthly cumulative Bering Strait ocean heat transport and the detrended monthly sea ice extent time series from May through September.

Results: An increase in the spring variability of the Bering Strait ocean heat transport is found since 2010, associated with both increased volume flux and water temperatures in May and June. A significant negative correlation between the Bering Strait ocean heat transport and Arctic sea ice extent in the Pacific sector is observed for May, June, and July, both within and outside the marginal ice zone, with a sharp decline in predictability for August and September.

Conclusion: The Bering Strait ocean heat transport is a skillful predictor for early melt season sea ice extent in the Pacific sector but loses predictive skills later in the summer in August and September due to changes in ice dynamics, in accordance with the loss of predictive skill in Global Climate Models.

Introduction

In recent decades, a sharp decline in sea ice extent (SIE), thickness, and age has been observed in the Arctic Ocean (1-3). These changes are projected to continue as per simulations using Coupled Model Intercomparison Project Phase 5 (CMIP5) under all representative concentration pathways (4). Increasing ocean heat is a significant contributor to sea ice decline (5). During the first large SIE decline of 2007, the heat transport through the Bering Strait was twice the 2001 heat flux, enough to account for approximately 30% of the estimated 2007 sea ice loss, and contributed to the creation of open water areas north of Bering Strait by May (6). Heat entering the broad shallow shelf of the Chukchi Sea (Fig. 1) interacts directly with the local sea-ice cover and can effectively reduce sea ice thickness before mixing with cooler waters from the central Arctic basin. Furthermore, dominant east-west sea ice drift in the winter and spring along the Alaskan coastline brings Pacific water heat from depth to the surface, due to coastal divergence and local Ekman offshore transport, leading to potentially thinner ice (7, 8). Additionally, loss of SIE and subsequent increases in the area of open water exposes ocean water to solar radiation directly, further warming the mixed layer and amplifying sea ice loss (i.e. the ice-albedo feedback) (9). While the volumetric transport through the Bering Strait is approximately 10% of the volume of the Fram Strait inflow, the yearly Bering Strait OHT (3 to 6×10^{20} Joules (J)) is the same order of magnitude as the Fram Strait OHT (5 to 13×10^{20} (J)) (5). The yearly OHT through the Barents Sea gate is significantly larger at approximately 2.21×10^{21} (J) (10). While outside the scope of this study, the transport of heat from the Atlantic Ocean via the Fram Strait and Barents Sea Gate plays a significant role in sea ice declines in the Atlantic Sector (11).

The Bering Strait, a narrow ~85-kilometer-wide channel between Alaska and Russia, is the only passage through which Pacific waters can enter the Arctic Ocean (5). The Bering Strait ocean heat transport (OHT) is a function of both water temperature and the volume transport of water through the strait. The large interannual variability of the Bering Strait OHT is thus the product of both the variability in the transport and temperature of the Bering Strait throughflow, which are themselves a function of both local and large-scale surface radiative and turbulent fluxes, surface winds, and internal oceanic variability (6). Quantifying the trends in Bering Strait OHT remains a difficult endeavor given the limited length of the Bering

Strait throughflow observational record, and currently, increases in annual water transport through the strait are the sole observed independent trends (5).

Previous studies have sought to better understand the influence of the Bering Strait OHT on sea ice in the Chukchi Sea at seasonal timescales. Woodgate et al. (2010) hypothesize that the Bering Strait OHT acts to weaken ice, precipitating the onset of solar driven sea ice melt, and providing a wintertime subsurface heat source due to large residence times of Pacific waters in the Arctic Ocean (6). Serreze et al. (2016) investigate the predictability of sea ice retreat and advance dates in the Chukchi Sea using the Bering Strait OHT (12). They find that the April-June throughflow accounts for 68% of retreat day variance, and that July-September throughflow accounts for 67% of advanced date variance (12). Additionally, they find a strong significant correlation ($r=0.8$) between the Bering Sea OHT from 1990 to 2013 (excluding 1993 to 1996 due to data gaps) and sea ice retreat date (12). While Serreze et al (2016) report on the seasonal response between the Bering Strait OHT and Chukchi sea ice thermodynamics, questions regarding the monthly response of Chukchi SIE to Bering Strait OHT variability within the melt season has not yet been investigated.

This study analyzes the variability and predictability of the Bering Strait OHT to better understand the sensitivity of Arctic Ocean SIE to Pacific Ocean heat fluxes at monthly timescales. We first assess the interannual and monthly variability of the Bering Strait OHT and then calculate the monthly, regional covariance of SIE and the Bering Strait OHT in the Chukchi and East Siberian Seas (ESS). Thirdly, we assess the spatial variability of the response of SIE across the Arctic Ocean in the context of melt season reductions in monthly maximum sea ice extents, and the marginal ice zone, where SIE is seasonally variable. The applications of these findings to seasonal sea ice forecasting are also discussed.

Methods

Bering Strait Ocean Heat Transport

The Bering Strait OHT was calculated using monthly averages of hourly, corrected, near-bottom temperature and transport observations from 1997 to 2015 from the A3 mooring (see Fig. 1b below), collected by Wood-

¹Editor Note: figures are adapted for print in black and white. View the coloured version online at msurj.com.

gate (2018) (5, 13). The dataset was post-processed in order to remove erroneous data and correct for sensor calibration errors (5). The A3 mooring is located approximately 35 kilometers north of the Bering Strait proper (Fig. 1b) and at a depth of 57 meters, roughly between 10 and 20 meters above the sea floor (5). This mooring was chosen based on the consistency and quality of the observational record and because its location in the outlet of the Bering Strait limits the influence of in-strait variations of velocity, temperature, or other parameters. Moorings A1 and A2, which are located in the Russian and US regions of the strait respectively, have temporally inconsistent and uncorrected records; and do not provide representative OHT estimates for the entire strait (5). Data from the A3 mooring, however, does not provide absolute OHT measurements due to its inability to account for the contribution of the Alaskan Coastal Current (ACC) (5). For the complete Bering Strait OHT, A3 recordings must be added to data from the A4 mooring, which was installed within the Strait off the Alaskan Coast in 2001 to better understand the role of the ACC in the Bering Strait throughflow (5). Due to the limited length of the A4 time series, it is not utilized in OHT calculations, resulting in the underestimation of the Bering Strait OHT. Nonetheless, the A3 mooring provides a sufficiently representative record to understand the relative variability of the throughflow and its influence on SIE in the Arctic Ocean (10). For more information of the mooring data used in this study, see Woodgate et al. (2018) and Woodgate (2015) (5, 13).

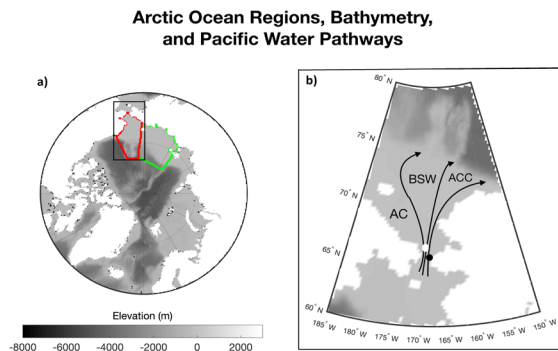


Figure 1. a) Bathymetry of Arctic Ocean and key regions. The Chukchi Sea is outlined in red and the East Siberian Sea is outline in green. Land areas are assigned a value of 3000 (m); b) pathways of Pacific waters (AC, Anadyr Current; BSW, Bering Shelf Water; and ACC, Alaskan Coastal Current) into the Chukchi Sea and Arctic Ocean. The approximate locations of the A3 (white) and A4 (black) moorings are shown. Plot (b) is adapted from Mathis et al. (2007) [19] (Note from the editor: figures are adapted for print in black and white. View the full coloured version online at msurj.com.)

The cumulative Bering Strait OHT, from Woodgate (2018), as assessed from January to month m , can be written as:

$$OHT_m = k \rho c_p \sum_{n=1}^m (T - T_{ref})_n V_n P_n$$

Where ρ is the density of ocean water assumed to constant (1023 kg/m^3), c_p is the specific heat of ocean water (3900 J/kgK), T is the monthly mean observed near-bottom temperature of water, which is considered representative of water column, in the Bering Strait in month n , T_{ref} is the reference freezing point temperature (-1.9°C) of ocean water at an approximate salinity of 32.5 psu. OHT is computed as relative to a reference freezing temperature at which Bering Strait waters leave the Arctic Ocean through the Fram Strait and Canadian Arctic Archipelago, thus allowing for the estimation of how much heat from Pacific waters has been lost to the Arctic Ocean throughout its transit (5). V_n (m^3/s) is the monthly mean volumetric transport of water through the Bering Strait during month n , P_n is the length of each month (s), and k (10^6) is a conversion factor (5). While the use of constant density, salinity, and reference temperature values introduces additional uncertainty to the OHT time series, Woodgate (2018) found that a salinity dependent reference freezing temperature would have relatively insignificant influence on the time series, only modifying OHT calculations by approximately 5% (5). Additionally, the use of a reference freezing point allows for a starting month of January, as win-

tertime months negligibly contribute to melt season OHT. The calculated Bering Strait OHT time series was detrended based on climatology before any statistical analyses were performed. For the purposes of examining the interannual variability of OHT, a monthly non-cumulative OHT time series was also created.

Sea Ice Extent

Arctic SIE is assessed using the NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration (14), dating back to November 1978. The dataset is based on brightness temperatures from the Special Sensor Microwave Imager/Sounder (SSMIS), the Special Sensor Microwave/Imager (SSM/I), and the Scanning Multichannel Microwave Radiometer (SMMR) to differentiate between sea ice and open water (14). Brightness temperatures are converted to sea ice concentration measurements using a rule-based classification (CDR algorithm) that relies on the highest output of two different proven algorithms, NASA Team and NASA Bootstrap (15). The dataset uses the Equal-Area Scalable Earth Grid (EASE-Grid), in which each grid cell has an area of 625 km^2 (15). We use the monthly resolution version of the NSIDC-CDR dataset.

SIE is defined as the sum of grid cells with greater than or equal to 15% sea ice coverage multiplied by the area of each grid cell. Grid cells with sea ice concentrations below 15% are considered to be open water. We define regions of the Chukchi and East Siberian Seas according to the NSIDC Arctic Ocean regional mask (see Fig. 1a above). Similar to OHT time series, SIE time series were detrended based on climatology before statistical analyses were conducted.

Results

Bering Strait Ocean Heat Transport Temporal Variability

Cumulative and non-cumulative heat transport through the Bering Strait was calculated from 1997 to 2015 (see Fig. 2-3). The non-cumulative OHT time series (see Fig. 3) shows that the use of a reference freezing temperature resulted in OHT values of 0 from January through April, which increase throughout the melt season from May to September, when the non-cumulative monthly transport is the largest, and then rapidly decreases during the fall. Interannual variability follows these seasonal cycles for both time series, increasing as the melt season progresses, with the least interannual variability throughout the melt season in May, and the most in September. As seen in Fig. 2, interannual variability has nearly doubled since 2010, with the range of observed cumulative September fluxes since 2010 at $2.31 \times 10^{20} \text{ Joules (J)}$ compared to $1.37 \times 10^{20} \text{ J}$ from 1997 – 2005. Additionally, the years 2007, 2011, 2015, which showed the largest September OHTs, were accompanied by anomalous OHT early in the summer (in May and June) unlike other years when the OHT anomalies starts in mid-summer in July. Percent change is used as an additional metric for Bering Strait OHT variability because of the magnitude scale differences between May and September heat fluxes. There is minimal variation in percent change throughout the year, with the exception of 2015, in which the May percent increase exceeded the September percent increase by ap-

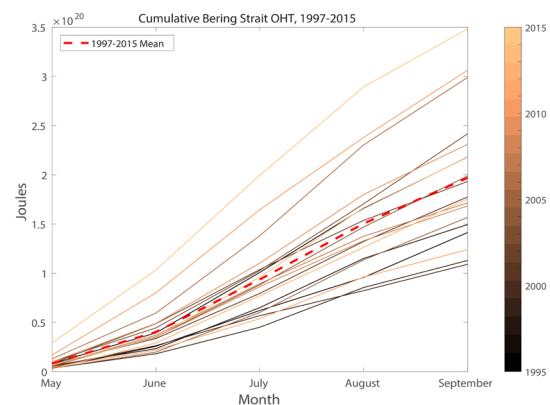


Figure 2. Cumulative monthly mean Bering Strait OHT (in Joules) from May of a given year through June-September of the same year. OHT values are relative to -1.9°C .

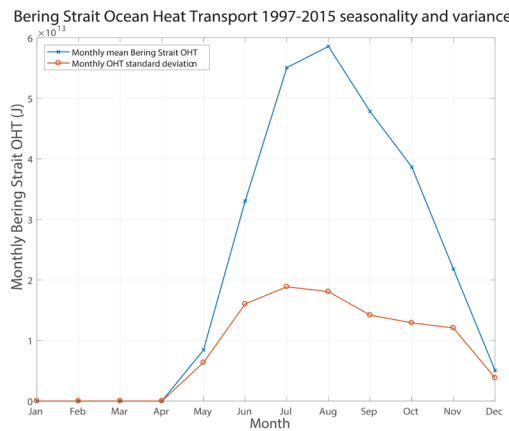


Figure 3. Monthly mean and standard deviation of Bering Strait OHT (in Joules) from 1997 to 2015. OHT values are relative to -1.9°C .

Regional Predictability

Month to month correlations were calculated between the Bering Strait OHT and SIE in the Chukchi Sea (see Table 1). Correlation coefficients above $r > |0.43|$ are statistically significant to the 95% confidence level using the standard t-test, and were found for the May, June, and July Bering Strait OHTs correlated with Chukchi SIEs over the same months. No significant correlations were found for August or September OHTs or SIEs. The largest significant correlations are seen in July SIEs with June and July Bering Strait OHTs which are correlated at $r \approx -0.8$, while the lowest significant correlation is observed between the May Bering Strait OHT and May Chukchi SIE at $r = -0.478$. Between May and July, the strength of the correlation increases as the melt season progresses. Month to month correlations were also calculated between the Bering Strait OHT and SIE in the East Siberian Sea (ESS) (see Table 1). Significant correlations are observed primarily in June SIE, with an insignificant positive correlation observed in May, and insignificant negative correlations in August and September. May OHTs were significantly correlated with July SIEs at exactly $r = 0.44$, while June and July OHTs were correlated with July SIEs at slightly below the 95% significance threshold.

Chukchi Sea	May SIE	June SIE	July SIE	August SIE	September SIE
May OHT	-0.48	-0.63	-0.71	-0.30	-0.17
June OHT	-	-0.70	-0.8	-0.35	-0.21
July OHT	-	-	-0.82	-0.36	-0.23
August OHT	-	-	-	-0.35	-0.25
September OHT	-	-	-	-	-0.24
ESS	May SIE	June SIE	July SIE	August SIE	September SIE
May OHT	0.23	-0.63	-0.44	-0.16	-0.08
June OHT	-	-0.70	-0.43	-0.17	-0.1
July OHT	-	-	-0.43	-0.20	-0.13
August OHT	-	-	-	-0.25	-0.19
September OHT	-	-	-	-	-0.17

Table 1. Correlation matrices between monthly Bering Strait OHT anomaly for month x and monthly mean Chukchi Sea and East Siberian Sea (ESS) SIE anomaly for month y . Correlations are 95% significant when $|r| > 0.43$. Significant correlations are shown in bold.

The sharp decline in late summer predictability is also seen in the anomaly persistence of the Bering Strait OHT and Chukchi SIE. Anomaly per-

sistence describes the influence of the variability of a climatologically detrended parameter later in its time series, and is calculated by correlating a given initial month with subsequent months of the same variable. Fig. 4 describes the anomaly persistence of the Bering Strait OHT and Chukchi SIE throughout the melt season. Bering Strait OHT anomalies have large significant persistence throughout the melt season from May through September with all intermonth correlations larger than 0.85. The persistence of Chukchi Sea SIE initialized in May shows a sharp continuous decline throughout the melt season, with significant persistence from May until July and insignificant persistence in August and September. Chukchi Sea SIE persistence for June and July shows a similar pattern, with a significant relationship for the first subsequent month, but then insignificant persistence in later months. Additionally, August Chukchi Sea SIE has high significant persistence with September SIE, with a correlation of $r = 0.83$ (not shown).

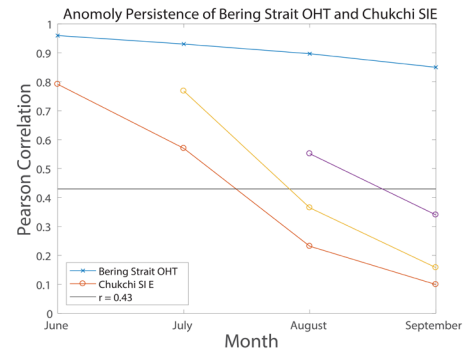


Figure 4. Anomaly persistence of the Bering Strait OHT (blue) and Chukchi SIE, from May to June through May to September (red), June to July through June to September (yellow), and July to August and September (purple). Correlations are 95% significant when $r > 0.43$ (black).

Spatial Correlations

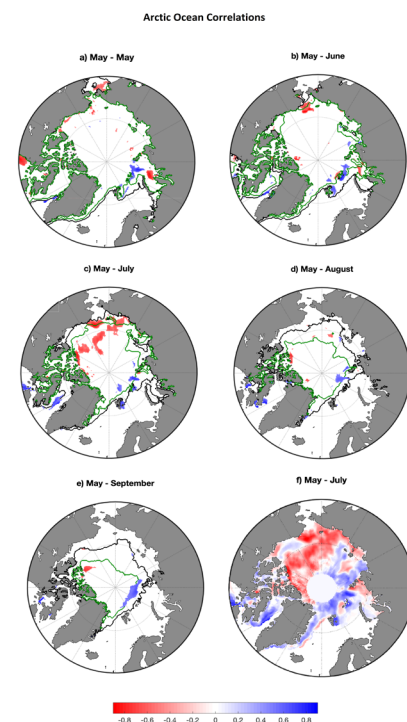


Figure 5. Significant Pearson correlations ($\geq 95\%$) across the Arctic Ocean for a) May Bering Strait OHT and May SIE, b) May Bering Strait OHT and June SIE, c) May Bering Strait OHT and July SIE, d) May Bering Strait OHT and August SIE, and e) May Bering Strait OHT and September SIE; f) shows all correlations between May Bering Strait OHT and July SIE. For a-e), the median maximum monthly SIE is outlined in black, and the minimum maximum monthly SIE is outlined in green.

Correlations between the Bering Strait OHT with the SIE were performed on a grid-cell basis in the Arctic Ocean (see Fig. 5). For May-May, May-June, and May-July correlations (Fig. 5a-c), regions of the Arctic Ocean with the highest correlations are in close proximity to the median ice edge. This means that in May, the region with the largest extent of significant negative correlations lies south of the Bering Strait; in June, the largest negative correlations are seen north of the mouth of the strait; and in July, further into the interior of the Chukchi Sea. Additionally, in July, there is an enlargement of the affected region in the Arctic Ocean to include regions of the ESS and the northern and eastern Beaufort Gyre. In Fig. 5d and 5e, the correlation maps of the May Bering Strait OHT and August and September SIE respectively, show an extreme reduction in regions affected by the Bering Strait OHT. For all months shown, affected regions include those both within and outside the marginal ice zone.

Positive correlations are observed in the Atlantic sector of the Arctic Ocean in areas near the median ice edge. For the May-May correlations (Fig. 5a), a positive signal is observed in the northern regions of the Barents and Kara Seas. For the May-June correlations (Fig. 5b), the signal is observed in the same regions but with a larger spatial extent. In the Barents and Kara Seas, the May-July (Fig. 5c) correlation shows a similar affected region to the May-May correlations, with the addition of significant positive correlations in the Baffin and Hudson Bays. Both the western and eastern regions of the Arctic Ocean see significant correlations with larger spatial extents when correlations are lagged, as seen in Fig. 5c. Additionally, a sub-seasonal oscillation is observed in the Hudson Bay, whereby, in May and June, a significant negative correlation is seen in the Western Hudson Bay, only to diminish in July and be replaced by significant positive correlations in Eastern Hudson Bay, which do not appear in May or June.

Discussion

May Bering Strait OHT is a skillful predictor for anomalously large September ocean heat fluxes, as seen in the large anomaly persistence throughout the entirety of the melt season in Fig. 5, and the large May and September OHTs in 2007, 2011, and 2015, as seen in Fig. 2-3. Woodgate (2018) finds that the anomalously large OHTs of 2007, 2011, and 2015 are due to the early arrival of warm waters by approximately 20 days, as well as rapid spring atmospheric warming (5). Increasing spring rapid warming leading to early sea ice breakup south of, and within, the Bering Strait, partially explains the increased OHT, and its associated variability. Sea ice breakup exposes ocean water to surface winds, increasing water velocity, which can increase transport into the Arctic, and thus increase variability.

The Bering Strait OHT has the greatest influence on Arctic SIE near the median monthly ice edge, which is expected given that the edges of the ice pack are most vulnerable to melt, especially from oceanic heat entering the Arctic Ocean from the south. This northward shift of the influence of the Bering Strait OHT throughout the melt season from the southern Bering Strait into the Chukchi Sea, and then further north into the western Arctic Ocean explains the increasing predictability trend from May through July. For May, the median monthly maximum SIE extends beyond the Bering Strait, thus a significant portion of the significant correlation is not included within the geographic limits of the Chukchi Sea. This is in contrast to June and July SIEs, in which the median maximum ice extent intersects the bounds of the Chukchi Sea, centering the OHT-affected regions within the Chukchi Sea limits, and thus producing a larger regional correlation. For August and September SIEs, the median ice edge is within the limits of the Chukchi Sea, but the region with significant correlations is reduced compared to earlier months, suggesting that the late summer decline in predictability is due to changes in sea ice dynamic and thermodynamic processes, such as drift (wind) and the ice-albedo feedback. Additionally, due to presence of significant correlations outside the marginal ice zone and within the minimum ice edges, we can conclude that the decline of predictability is due to dynamics and not due to the regions becoming ice-free. We hypothesize that significant positive correlations in the Atlantic sector of the Arctic Ocean are caused by increases in the volume flux through the Bering Strait leading to decreased OHT through the Fram Strait and Barents Sea Gate, and thus increased SIE, due to the conservation of water mass of the Arctic Ocean, as seen in Jahn et al. (2010) (16). These conclusions are also supported by Auclair & Tremblay (2018), who utilized the Community Earth System Model Large Ensemble (CESM-LE)

to test the relationship between different OHTs and rapid sea ice declines, in which absorbed shortwave radiation by the ocean from May through September was significantly, negatively correlated with Bering Strait OHT anomalies in the Barents and Kara Seas, also suggesting a positive relationship between the Bering Strait OHT and SIE in the Atlantic sector of the Arctic Ocean (11).

We find that May Bering Strait OHT is a poor predictor for the end of summer (August and September) SIE in the Chukchi Sea (see Table 1). These conclusions are also supported by Auclair & Tremblay (2018), who posit that OHTs had the greatest influence on SIE over shallow continental shelves, in which the majority of heat remains in the mixed layer, whereas in the Arctic basin, vertical heat transport is reliant on Ekman pumping, thus decreasing the exposure of retreated pack ice to oceanic heat (11). They also found decreased coverage of significant negative correlations between Bering Strait OHT anomalies and Chukchi SIE, in agreement with our findings (11). The diminished influence of Bering Strait OHT on late summer SIE is in accord with the idea of a predictability barrier for the summer SIE, as investigated in model-based forecast studies by Day et al. (2014) and Bushuk et al. (2017) (17, 18). Using perfect-model experiments based on the Geophysical Fluid Dynamics Laboratory seasonal prediction system, Bushuk et al. found that predictions of September SIE in the Chukchi Sea are skillful up to a lead time of two months – i.e.: there is a sharp loss of skill for forecasts initialized before July, hence a ‘July barrier for predictability’ (18). Our findings of a lack of correlation between early summer OHT and late summer SIE support the idea that late summer thermodynamic and dynamic processes, such as surface winds, air temperatures, and open water exposed to direct solar heating, drive the variability of summer SIE in the Chukchi Sea, overriding any early-season signature from ocean heat (9). Our results and hypothesized mechanisms, however, also suggest a significant role that the Bering Strait OHT plays on early-summer SIE (up to July), in agreement with Serreze et al. (2016), who note that the April to June oceanic heat inflow through the Bering Strait is strongly correlated with the retreat date of sea ice in the Chukchi Sea (12). Anomalies in other parameters such as sea ice thickness may have longer memory, enabling skillful September sea ice predictions from as early as May (18). Additionally, our findings illustrate the inherent complexity of predicting SIE via global climate models, in which a multitude of parameters gain and lose predictive ability throughout the year.

More work is required in order to adequately explain several of the results of this study. Firstly, there is a discrepancy between the ESS regional correlations and the July spatial correlations, outlined in Table 1 and Fig. 5 respectively. The spatial extent of the negative signal in the ESS in Fig. 5c, would suggest significant regional correlations in July in addition to June. This discrepancy is the result of insignificant positive correlations in other regions of the ESS as seen in Fig. 5f, which partially oppose the negative correlations enough to bring the regional signal to below the 95% significance level. More work, however, is required in order to understand the source of this positive signal in the ESS in July. Secondly, Serreze et al. (2016), found a significant correlation (~ 0.67) between July through September Bering Strait OHT and sea ice advance date in the Chukchi Sea, whereas we found no significant correlations between the May through September Bering Strait OHT and September sea ice extent (12). This predictability disparity is unexpected given the similarity between advanced date, first date of the year when sea ice concentration exceeds 30%, and SIE, as well as the agreement between early melt season retreat date and SIE predictabilities (12). Additionally, the high persistence of Bering Strait OHT anomalies suggests that it is highly unlikely that the discrepancies in heat flux start date (July vs. May) are responsible for the large differences in predictability. This discrepancy is likely due to different regional definitions of the Chukchi Sea, which in Serreze et al. (2016) is not defined beyond the broad-shallow shelf, whereas our boundaries extend into the deeper waters of the Arctic Ocean interior (12). More research on the differing mechanisms governing Chukchi Sea advanced date and SIE beyond the Chukchi Sea shelf is required in order to understand this disagreement.

Conclusion

This study investigated the relationship between the Bering Strait ocean heat transport (OHT) and sea ice extent (SIE) in the western Arctic Ocean at monthly time scales, as well as patterns in variability and predictability

across the Bering Strait OHT time series from 1997 to 2015. We found increased variability in the Bering Strait OHT since 2010, with large heat flux anomalies early in the melt season in May and June as strong predictors for large September OHTs. Strong negative correlations were observed between Bering Strait OHT and SIE in the Chukchi Sea during May, June, and July, with a sharp decline in predictability for August and September. High predictability was also found between May Bering Strait OHT flux and July SIE in the Northern and Eastern Beaufort Gyre, as well as along the Siberian Coast and parts of the East Siberian Sea. These results not only contribute to our understanding of ice-ocean dynamics in the Arctic Ocean, but also provide observational evidence in support of SIE predictability for both seasonal forecasts and global climate models. Improved sea ice prediction will help ensure the safe functioning of industrial and shipping operations in the Arctic Ocean as the Arctic takes on greater global economic importance.

Acknowledgements

We would like to thank the McGill School of Environment and Department of Atmospheric and Oceanic Sciences for their support throughout this research, and our three anonymous reviewers for their constructive feedback.

References

1. Stroeve J, Holland MM, Meier W, Scambos T, Serreze M. Arctic sea ice decline: Faster than forecast. *Geophysical Research Letters*. 2007Jan;34(9).
2. Kwok R, Rothrock DA. Decline in Arctic sea ice thickness from submarine and ICESat records: 1958–2008. *Geophysical Research Letters*. 2009 Jun;36(15).
3. Maslanik JA, Fowler C, Stroeve J, Drobot S, Zwally J, Yi D, et al. A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss. *Geophysical Research Letters*. 2007;34(24).
4. Stroeve J, Notz D. Changing state of Arctic sea ice across all seasons. *Environmental Research Letters*. 2018;13(10):103001.
5. Woodgate RA. Increases in the Pacific inflow to the Arctic from 1990 to 2015, and insights into seasonal trends and driving mechanisms from year-round Bering Strait mooring data. *Progress in Oceanography*. 2018;160:124–54.
6. Woodgate RA, Weingartner T, Lindsay R. The 2007 Bering Strait oceanic heat flux and anomalous Arctic sea-ice retreat. *Geophysical Research Letters*. 2010;37(1).
7. Holland MM, Bitz CM, Tremblay B. Future abrupt reductions in the summer Arctic sea ice. *Geophysical Research Letters*. 2006 Dec;33(23).
8. Pickart RS, Moore G, Mao C, Bahr F, Nobre C, Weingartner TJ. Circulation of winter water on the Chukchi shelf in early Summer. *Deep Sea Research Part II: Topical Studies in Oceanography*. 2016;130:56–75.
9. Perovich DK, Light B, Eicken H, Jones KF, Runciman K, Nghiem SV. Increasing solar heating of the Arctic Ocean and adjacent seas, 1979–2005: Attribution and role in the ice-albedo feedback. *Geophysical Research Letters*. 2007 Nov;34(19).
10. Smedsrud LH, Esau I, Ingvaldsen RB, Eldevik T, Haugan PM, Li C, et al. The Role of The Barents Sea In The Arctic Climate System. *Reviews of Geophysics*. 2013;51(3):415–49.
11. Auclair G, Tremblay LB. The Role of Ocean Heat Transport in Rapid Sea Ice Declines in the Community Earth System Model Large Ensemble. *Journal of Geophysical Research: Oceans*. 2018 Nov;123: 8941–8957
12. Serreze MC, Crawford AD, Stroeve JC, Barrett AP, Woodgate RA. Variability, trends, and predictability of seasonal sea ice retreat and advance in the Chukchi Sea. *Journal of Geophysical Research: Oceans*. 2016;121(10):7308–25.
13. Woodgate R, Stafford K, Prah F. A Synthesis of Year-Round Interdisciplinary Mooring Measurements in the Bering Strait (1990–2014) and the RUSALCA Years (2004–2011). *Oceanography*. 2015Jan;28(3):46–67. Available from: <http://psc.apl.washington.edu/HLD/Bstrait/Data/Bering-StraitCruiseDataArchive.html>
14. Meier W, Fetterer F, Savoie M, Mallory S, Duerr R, et al. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3. 2017; NSIDC: National Snow and Ice Data Center.
15. Peng G, Meier WN, Scott DJ, Savoie MH. A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring. *Earth System Science Data Discussions*. 2013Jul;6(1):95–117.
16. Jahn A, Tremblay B, Mysak LA, Newton R. Effect of the large-scale atmospheric circulation on the variability of the Arctic Ocean freshwater export. *Climate Dynamics*. 2009;34(2–3):201–22.
17. Day JJ, Tietsche S, Hawkins E. Pan-Arctic and Regional Sea Ice Predictability: Initialization Month Dependence. *Journal of Climate*. 2014;27(12):4371–90.
18. Bushuk M, Msadek R, Winton M, Vecchi GA, Gudgel R, Rosati A, et al. Skillful regional prediction of Arctic sea ice on seasonal timescales. *Geophysical Research Letters*. 2017;44(10):4953–64.
19. Mathis JT, Pickart RS, Hansell DA, Kadko D, Bates NR. Eddy transport of organic carbon and nutrients from the Chukchi Shelf: Impact on the upper halocline of the western Arctic Ocean. *Journal of Geophysical Research*. 2007 Apr;112(C5).

Research Article

¹McGill University, Montreal,
QC, Canada

Keywords

Evolution, Anolis, ecomorphs,
traits, predation, ecology

Email Correspondence

samantha.wunderlich@mail.
mcgill.ca

Samantha Wunderlich¹

Predator Driven Trait Changes in *Anolis smaragdinus*

Abstract

Background: Research on the evolution of traits in regard to predator-prey relationships can illuminate the effects of migrating species and invasive species on native populations, which is increasingly relevant with heightened anthropogenic environmental disturbances shifting territories of native species.

Methods: We used a replicated field experiment to test the effects that predators have on the phenotype of *Anolis smaragdinus* in the Bahamas. It was predicted that *A. smaragdinus* would rapidly evolve over six generations to have shorter fore and hind limb lengths due to the presence of the predator, *Leiocephalus carinatus*, driving *A. sagrei* and *A. smaragdinus* into higher habitats, which selects for shorter limb lengths.

Results: ANCOVAs were performed for male and female *A. smaragdinus*, where only two of the thirteen measured traits, ulna length and third toe length in females, were found to have a significant year by treatment interaction between 2011 and 2017. We observed a general reduction in mean trait size when predators are present. Therefore, predators may be driving *A. smaragdinus* to higher perching sites which selects for smaller limb length. Also, it was found that populations of *A. smaragdinus* may be able to evolve in as little as six generations when predators and competitors are present.

Limitations: A longer study is needed to see if other trait changes would become significant over time, since the residual means are trending in one direction. Another suggestion is to do a study with fewer experimental islands and larger founder populations. This would violate the meaning of a true founder population; however, it may allow *A. smaragdinus* to become fully integrated within the new environment and lead to a smaller chance of extinction.

Conclusion: The study showed that the introduction of predators may cause significant phenotypic changes in female *A. smaragdinus*' ulna and third toe length, while also showing a general reduction in mean trait size when predators are present. Therefore, predators may be driving *A. smaragdinus* to higher perching sites which selects for smaller limb length. However, the study did not show a significant change in either the fore limb length or the hind limb length between 2011 and 2017 which was predicted originally.

Introduction

With a rapidly changing climate, species are being driven into predator-occupied territories to remain in optimal environmental conditions (1). It is essential to study how species interact with each other when they have previously been ecologically naive, since the interactions between invasive and native species can affect each group's behavior, morphology, and survival rates (2). Research on trait evolution resulting from predator-prey relationships can illuminate the effects of migrating species and invasive species on native populations, which is increasingly relevant with heightened anthropogenic environmental disturbances.

There are many examples of directional trait changes in prey species, such as the directional evolution of the peppered moth towards more darkly-colored wings in the presence of high smog levels (3-4) and the evolution of Mullerian mimicry in butterflies (5-6). These classic studies exemplify a directional evolutionary response caused by increased predation on organisms that are not adapted to a novel or changing environment. These are just some of the possible mechanisms that prey species have evolved to escape predation; another mechanism is behavioral change. It has previously been shown that when predators are introduced to prey, the latter change their use of habitat (7-8). This mechanism was illustrated by studies showing crayfish moving to more protective habitats in response to the presence of predators (9-10). Therefore, prey are able to adjust their use of habitat as a strategy to avoid predation.

Anolis is a genus that has been studied extensively with regard to mor-

phology and behavior in the Bahamas (11). *Anolis* species have repeatedly radiated into different ecomorphs, which are adapted to different microhabitats (15-17). The species *Anolis smaragdinus* and *Anolis sagrei* are two different ecomorphs which have developed said adaptations (18). *A. sagrei* is a competitive species that is able to reach high population densities, is highly invasive, and can outcompete many native species of lizards (19-21). It is part of the trunk-ground ecomorph, signifying that it perches on the lower trunks of trees and on objects on the ground. *A. smaragdinus* is part of the trunk-crown ecomorph, as it has a tendency to occupy the uppermost trunks of tall trees and the lower canopy. Consequently, it is better adapted to perching and running on smaller branches (22) and it is larger than *A. sagrei* (23).

Previous research has found that *A. sagrei* moves to higher vegetation to escape the predator, *Leiocephalus carinatus*, (24). Similarly, *A. carolinensis* will partition its habitat when *A. sagrei* is introduced as a competitor species (25). *A. smaragdinus* is subgroup of *A. carolinensis* (26); therefore, they are in the same ecomorph and are highly related. Shorter limbs of *A. carolinensis* and *A. smaragdinus* help increase manoeuvrability on branches; previous research has found that in *A. sagrei* natural selection favours shorter hind limbs when they occupy higher perches (27). However, that study did not look at the changes in limb length over multiple generations. Another study has found that the toe pads of *A. carolinensis* evolve rapidly when *A. sagrei* are introduced (28). Literature regarding predator introduction and the consequent habitat shift of *A. sagrei* on other species of lizard, such as *A. smaragdinus*, is lacking. It has been suggested that two closely co-occurring related species could diverge when interactions

are negatively occurring (29). Therefore, we predict that *A. smaragdinus* will rapidly evolve over six generations to have shorter fore and hind limb lengths due to the presence of the predator, *L. carinatus*, driving *A. sagrei* and *A. smaragdinus* into higher habitats, which selects for shorter limb lengths due to competition. Additionally, by using a portable x-ray machine to measure changes in *A. smaragdinus*' traits between 2011 and 2017, we will be able to obtain more precise measurements compared to using a tape measure or a caliper since gravidity can affect growth (30).

Methods

Field Site

16 small islands located near Staniel Cay, in Exuma Bay of Grand Bahama Bank (26.17°, -76.44°) were chosen for our experiment. These islands have been used for experimentation since 1977 (31). Islands were included in the study if they met the following criteria: presence of robust *Anolis sagrei* (brown anole) population; absence of any other detectable lizard or top-predator species; and presence of trees more than two meters tall which are deemed necessary to support *Anolis smaragdinus* (green-anole) populations (32). *A. sagrei* was censused on the 16 islands in May 2011, during which time no other lizard species were detected.

Islands were given one of four treatments: control islands had only *A. sagrei* on them as the native lizard species, and the remaining experimental islands had *A. sagrei* with *L. carinatus* and *A. smaragdinus*; *A. sagrei* with only *A. smaragdinus*; or *A. sagrei* with only *L. carinatus*. Island treatments were assigned by a random coin flip. Islands were first stratified by size and measured in terms of vegetated area, such that two replications of each treatment would be assigned to the eight smallest and the eight largest islands. *A. smaragdinus* were collected from Staniel Cay by noose since *L. carinatus* were absent on this island; thereby both *Anolis* were ecologically naive to the predator, *L. carinatus*. On other islands these lizards co-exist, therefore they were not evolutionarily naive.

Field Work

In 2011, experimental islands were surveyed for initial lizard species and populations that were present. *A. smaragdinus* was caught opportunistically from Staniel Cay with a noose and pole or by hand and kept captive for a couple of days. We used Isoflurane to anesthetize the lizards (33) and x-rayed them using a portable x-ray machine. After x-rays were taken, the lizards were weighed, and a small piece of tail tissue was obtained for DNA sequencing. *A. smaragdinus* was assigned randomly in groups of ten or eleven to islands with five to seven *L. carinatus* of unknown sex. These numbers were chosen to simulate colonization by small founder populations (34).

In 2017, we surveyed the lizard populations on the experimental islands and caught *A. smaragdinus* to measure phenotype traits with a portable x-ray scanner, holding them captive for a maximum of 24 hours. After x-rays were taken, the lizards were weighed and a small piece of tail tissue was extracted for DNA sequencing. In order to prevent the immediate recapture of individuals, each lizard was marked with nail polish that would be removed with the lizards next molting. When catching lizards, bodies were checked for nail polish and tail integrity to ensure no recapture of individuals. *A. smaragdinus* were found on only two of the four experimental islands at the time of measurement.

Analysis

One researcher measured skeletal traits related to locomotion and niche using the ObjectJ plugin for ImageJ (35; Fig. 1). Measurements were taken on both sides of the *A. smaragdinus* sagittal plane in order to average the measurements. If a lizard had an abnormality, such as a missing claw, then only the data from the undamaged side of the lizard was used for that specific measurement.

R (v. 1.1.456) (36) was used to prepare the data for analysis. Sagittal plane measurements that were averaged were converted into millimeters. *A.*

smaragdinus were checked for recapture by looking at the Euclidean distance between measurements of each island by sex, and if this value was zero then the x-rays and field notebook data was analyzed. Additionally, only adult lizards (male snout-vent length (SVL) > 45.0 mm; 37) were used in subsequent analyses. Due to unclear literature regarding size cut-offs for adult females, a threshold was obtained using graphical analysis which determined a cut off of SVL > 37.5 mm. After data cleaning, there were a total of 39 lizards removed based on SVL measurements and field data. Analysis of Covariance (ANCOVA) was used to evaluate changes in phenotype measurements over time with respect to experimental treatment. Since *A. smaragdinus* are sexually dimorphic (38), analysis was separated by sex. We pooled lizards by treatment, whether a *L. carinatus* was present or absent, and used linear models of the form trait ~ SVL + year*treatment with the uncorrected size measurements. The year by treatment interaction was used to test whether predation caused phenotypic evolution.

In order to examine the phenotypic trait changes more generally, we looked at the changes in the relative size-corrected traits through time by treatment. Size corrections were performed using SVL as a covariate for body size where the relative trait measurements were the residuals from sex-specific linear regressions of trait by SVL. These measurements were compared to the residual lengths means of each trait to determine the direction of trait change found in the ANCOVA.

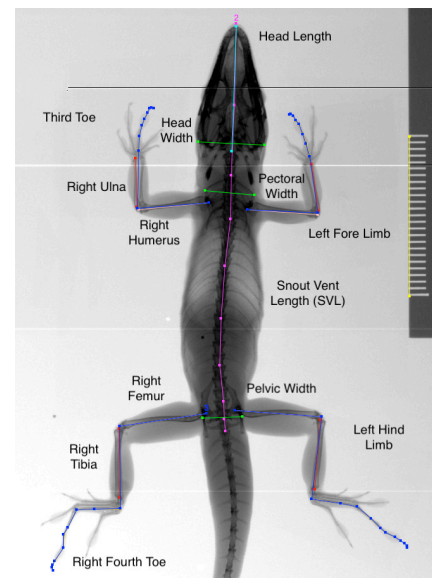


Fig. 1. Object J Screenshot of Vector Measurements of Skeletal Traits. Measured traits include snout-vent-length (SVL), head width, head length, pectoral width, right and left humerus, right and left ulna, right and left third toe, right and left forelimb, pelvic width, right and left femur, right and left tibia, right and left fourth toe, right and left hind limb, and ruler.

Results

In 2011, we analyzed 78 lizards and in 2017, we analyzed 83 lizards; 84 female and 77 male (N=161) x-rays of *A. smaragdinus* were used for this study. For both males and females, measured traits tended to become smaller between 2011 and 2017 on islands with predators and larger on islands without predators. When predators were present on the islands, the means of the residual lengths of the traits decreased from 2011 to 2017 in all traits for the males except their pelvic width, humerus, and SVL (which all increased). When the predators were absent, the means of the residual lengths increased in all traits except for head width, pelvic width, and femur length, where they decreased. In females the means of the residual lengths of traits decreased from 2011 to 2017 for all traits except for pelvic width and pectoral width when predators were present. When predators were absent, all the means of the trait residual lengths increased in females.

We performed ANCOVAs for male *A. smaragdinus*, where none of the thirteen measurable traits were found to have a significant year by treat

ment interaction between 2011 and 2017 ($p > 0.05$) when using the uncorrected traits. ANCOVAs were conducted for female *A. smaragdinus* to determine if there was variation between the trait measurements on a year by treatment interaction between 2011 and 2017. Only two of thirteen measured traits were found to have a significant year by treatment interaction: third toe length ($F_{1,79} = 4.292$, $p = 0.0415$; Fig. 2) and ulna length ($F_{1,79} = 4.231$, $p = 0.043$; Fig. 2). The rest of traits were not found to be significant on a year by treatment interaction basis ($p > 0.05$).

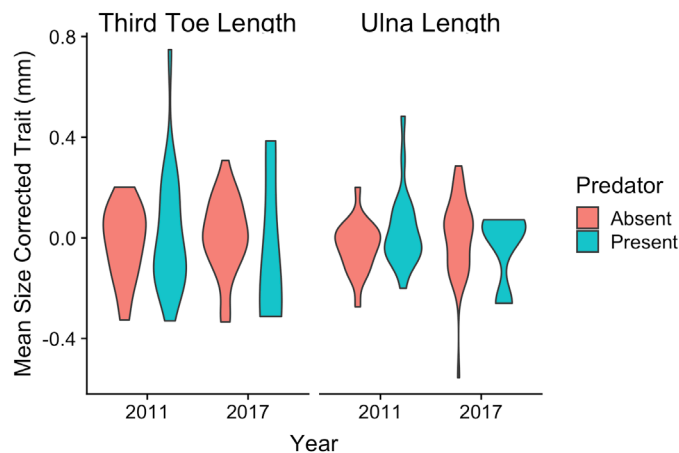


Fig. 2. Violin Plot of the Mean Size Corrected Traits, Third Toe Length and Ulna Length. This plot looks at the change in the mean size corrected trait in millimeters between years 2011 and 2017 for females on islands with or without predators.

Discussion

We used a replicated field experiment to test the effects that predators have on *Anolis smaragdinus* phenotype in the Bahamas. It was predicted that *A. smaragdinus* would rapidly evolve over six generations to have shorter fore and hind limb lengths due to the presence of the predator, *L. carinatus*, driving *A. sagrei* and *A. smaragdinus* into higher habitats, which selects for shorter limb lengths. ANCOVAs were performed for male and female *A. smaragdinus*. Only two of the thirteen measured traits were found to have a significant year by treatment interaction between 2011 and 2017: ulna length and third toe length in females.

In general, mean residual lengths of the traits decreased about one millimeter from 2011 to 2017 when predators were present. Notable exceptions are SVL in males and pelvic width and pectoral width in females, where the mean residual lengths increased. When predators were absent the traits tended to increase in their mean residual lengths. However, in males none of these differences were significant, and in females only the changes in the third toe length and ulna length were significant on $p < 0.05$ in the ANCOVA; we did not find a significant change in the whole fore limb or the full hind limb in the ANCOVA. A possible explanation for the reduction in limb length trends is that when a predator is introduced into *A. smaragdinus*' habitat, the lizards change the way they use their habitats. This would coincide with previous research in that lizards will begin to perch higher to avoid the predator (39). Another possible reason for the reduction in the mean residual trait lengths between 2011 and 2017 is that islands are generally populated by a small population of colonizers that are especially susceptible to genetic drift. Over a period of six generations, it is possible for extremely small founder populations to experience genetic drift due to the founder effect. However, previous research has found that *Anolis* maintains genetic diversity over seven generations in small populations on Dominica (40). Also, the mean residual lengths of the traits tended to increase in predator absent islands. This may be explained by the lack of predators preying on *A. sagrei* and *A. smaragdinus*, which usually forces *A. smaragdinus* into higher vegetation and selects for smaller phenotypes (41). Furthermore, *A. smaragdinus* were transplanted from

very large islands with very tall vegetation to smaller islands. Therefore, the changes we see may be due to adapting to new, smaller islands rather than the predator's effects. Previous research has found that natural selection can adjust morphology of lizards after exposure to habitats of different structures (42).

We may not have found significance in many traits in the ANCOVA due to the low population size of *A. smaragdinus* and high population densities of *A. sagrei* in our experiment. In order to simulate a founder population, we only transplanted 10-11 *A. smaragdinus* on to each island, simulating a colonization event. This gives a low statistical power for our ANCOVA. Additionally, only two out of four experimental islands had *A. smaragdinus* remaining on them. Therefore, our results are conditional that the lizards survive to establish a colony from 2011 and 2017. Furthermore, all experimental islands had high density populations of *A. sagrei*. Even though they are less adapted to arboreal habitats, their sheer density may have led to interspecific competition between them and *A. smaragdinus* for habitat space, leading to them outcompete *A. smaragdinus*.

Our research has shown that phenotypic traits can change significantly in as little as six generations when strong enough selection is present. This coincides with previous research showing that the interaction of species can cause phenotypic changes granted that the selection is strong enough on the trait (43). However, another study found that when birds and snakes were introduced to *A. sagrei* populations, selection did not alter phenotypic traits (44). Therefore, different types of predators may have different effects on prey and competitive species. Additionally, a longer study is needed to see if other trait changes would become significant over time, since the residual means are trending in one direction. Another suggestion is to do a study with fewer experimental islands and larger founder populations. This would violate the idea of a true founder population being a small group of members; however, it may be relevant as populations of species are migrating or invasive species are introduced into new habitats. Furthermore, it would ensure success of the *A. smaragdinus* and it would limit confounding variables thereby increasing colonization success (45). One technical advantage of our study is that we used x-ray measurements to compare the differences in limb lengths between 2011 and 2017. This gave us more precise measurements than measuring the outer limbs with a measuring tape or caliper since the bones themselves were measured and not the bones and tissue. Lastly, this allowed us to measure traits not frequently looked at like pectoral width and pelvic width.

Conclusion

Our study shows that the introduction of predators may cause significant phenotypic changes in female *A. smaragdinus*' ulna and third toe length. It also shows a general reduction in mean trait size when predators are present. Therefore, predators may be driving *A. smaragdinus* to higher perching sites which selects for smaller limb length. Our study did not show a significant change in either the fore limb length or the hind limb length between 2011 and 2017. Therefore, we show that populations of *A. smaragdinus* may be able to evolve in as little as six generations when predators and competitors are present. This is relevant since anthropogenic environmental disturbances may push invasive species and predators into new territories, therefore affecting the morphology of other species.

Acknowledgements

I would like to express my deepest appreciation to my supervisors, Timothy Thurman and Dr. Rowan Barrett, for allowing me to analyze a chunk of their data and guiding me in the process. I appreciate all the valuable skills I was able to learn while I was undertaking this project.

References

1. Rahel FJ, Olden JD. Assessing the effects of climate change on aquatic invasive species. *Conservation biology*. 2008;22(3):521-33.
2. Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE. The spread of invasive species and infectious disease as drivers of ecosystem change. *Frontiers in Ecology and the Environment*. 2008;6(5):238-46.

3. Clarke CA, Mani G, Wynne G. Evolution in reverse: clean air and the peppered moth. *Biological Journal of the Linnean Society*. 1985;26(2):189-99.
4. Grant BS. Fine tuning the peppered moth paradigm. *Evolution*. 1999;53(3):980-4.
5. Turner JR. Studies of Müllerian mimicry and its evolution in burnet moths and heliconid butterflies. *Ecological genetics and evolution*: Springer; 1971. p. 224-60.
6. Kapan DD. Three-butterfly system provides a field test of Müllerian mimicry. *Nature*. 2001;409(6818):338.
7. Schoener TW, Spiller DA, Losos JB. Predators increase the risk of catastrophic extinction of prey populations. *Nature*. 2001;412(6843):183.
8. Schoener TW, Spiller DA, Losos JB. Predation on a common *Anolis* lizard: can the food-web effects of a devastating predator be reversed? *Ecological Monographs*. 2002;72(3):383-407.
9. Stein RA, Magnuson JJ. Behavioral response of crayfish to a fish predator. *Ecology*. 1976;57(4):751-61.
10. Mittelbach G. Predator-mediated habitat use: some consequences for species interactions. *Environmental Biology of Fishes*. 1986;16(1-3):159-69.
11. Johnson MA, Revell LJ, Losos JB. Behavioral convergence and adaptive radiation: effects of habitat use on territorial behavior in *Anolis* lizards. *Evolution: International Journal of Organic Evolution*. 2010;64(4):1151-9.
12. Andrews RM. Growth rate in island and mainland anoline lizards. *Copeia*. 1976:477-82.
13. Martin AP, Palumbi SR. Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences*. 1993;90(9):4087-91.
14. Reznick DN, Shaw FH, Rodd FH, Shaw RG. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*. 1997;275(5308):1934-7.
15. Losos JB, Jackman TR, Larson A, de Queiroz K, Rodriguez-Schettino L. Contingency and determinism in replicated adaptive radiations of island lizards. *Science*. 1998;279(5359):2115-8.
16. Schluter D. *The ecology of adaptive radiation*: OUP Oxford; 2000.
17. Butler MA, Sawyer SA, Losos JB. Sexual dimorphism and adaptive radiation in *Anolis* lizards. *Nature*. 2007;447(7141):202.
18. Losos JB. Ecomorphology, performance capability, and scaling of West Indian *Anolis* lizards: an evolutionary analysis. *Ecological Monographs*. 1990;60(3):369-88.
19. Losos JB, Marks JC, Schoener TW. Habitat use and ecological interactions of an introduced and a native species of *Anolis* lizard on Grand Cayman, with a review of the outcomes of anole introductions. *Oecologia*. 1993;95(4):525-32.
20. Campbell TS. Analyses of the effects of an exotic lizard (*Anolis sagrei*) on a native lizard (*Anolis carolinensis*) in Florida, using islands as experimental units. 2000.
21. Gerber GP, Echternacht AC. Evidence for asymmetrical intraguild predation between native and introduced *Anolis* lizards. *Oecologia*. 2000;124(4):599-607.
22. Losos JB, Irschick DJ. The effect of perch diameter on escape behaviour of *Anolis* lizards: laboratory predictions and field tests. *Animal Behaviour*. 1996;51(3):593-602.
23. Jenssen TA, Congdon JD, Fischer RU, Estes R, Kling D, Edmands S. Morphological characteristics of the lizard *Anolis carolinensis* from South Carolina. *Herpetologica*. 1995:401-11.
24. Schoener TW, Spiller DA, Losos JB. Predation on a common *Anolis* lizard: can the food-web effects of a devastating predator be reversed? *Ecological Monographs*. 2002;72(3):383-407.
25. Kamath A, Stuart YE, Campbell TS. Behavioral Partitioning by the Native Lizard *Anolis carolinensis* in the Presence and Absence of the Invasive *Anolis sagrei* in Florida. *Breviora*. 2013:1-10.
26. Hedges SB, Burnell KL. The Jamaican radiation of *Anolis* (Sauria: Iguanidae): an analysis of relationships and biogeography using sequential electrophoresis. *Caribbean Journal of Science*. 1990;26(1-2):31-44.
27. Losos JB, Schoener TW, Langerhans RB, Spiller DA. Rapid temporal reversal in predator-driven natural selection. *Science*. 2006;314(5802):1111-.
28. Stuart YE, Campbell T, Hohenlohe P, Reynolds RG, Revell L, Losos J. Rapid evolution of a native species following invasion by a congener. *Science*. 2014;346(6208):463-6.
29. Brown WL, Wilson EO. Character displacement. *Systematic zoology*. 1956;5(2):49-64.
30. Cox RM, Calsbeek R. Severe costs of reproduction persist in *Anolis* lizards despite the evolution of a single-egg clutch. *Evolution*. 2010;64(5):1321-30.
31. Schoener TW, Schoener A. Distribution of vertebrates on some very small islands. I. Occurrence sequences of individual species. *The Journal of Animal Ecology*. 1983:209-35.
32. Losos JB, Spiller DA. Differential colonization success and asymmetrical interactions between two lizard species. *Ecology*. 1999;80(1):252-8.
33. Bennett RA, editor *Reptile anesthesia*. Seminars in avian and exotic pet medicine; 1998: Elsevier.
34. Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB. Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science*. 2012;335(6072):1086-9.
35. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*. 2012;9(7):676.
36. Team RC. R: A Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.
37. Jenssen TA, Congdon JD, Fischer RU, Estes R, Kling D, Edmands S. Morphological characteristics of the lizard *Anolis carolinensis* from South Carolina. *Herpetologica*. 1995:401-11.
38. Butler MA, Sawyer SA, Losos JB. Sexual dimorphism and adaptive radiation in *Anolis* lizards. *Nature*. 2007;447(7141):202.
39. Losos JB, Schoener TW, Langerhans RB, Spiller DA. Rapid temporal reversal in predator-driven natural selection. *Science*. 2006;314(5802):1111-.
40. Eales J, Thorpe R, Malhotra A. Weak founder effect signal in a recent introduction of Caribbean *Anolis*. *Molecular Ecology*. 2008;17(6):1416-26.
41. Losos JB, Schoener TW, Langerhans RB, Spiller DA. Rapid temporal reversal in predator-driven natural selection. *Science*. 2006;314(5802):1111-.
42. Vitt LJ, Caldwell JP, Zani PA, Titus TA. The role of habitat shift in the evolution of lizard morphology: evidence from tropical *Tropidurus*. *Proceedings of the National Academy of Sciences*. 1997;94(8):3828-32.
43. Brown WL, Wilson EO. Character displacement. *Systematic zoology*. 1956;5(2):49-64.
44. Calsbeek R, Cox RM. Experimentally assessing the relative importance of predation and competition as agents of selection. *Nature*. 2010;465(7298):613.
45. Ahlroth P, Alatalo RV, Holopainen A, Kumpulainen T, Suhonen J. Founder population size and number of source populations enhance colonization success in waterstriders. *Oecologia*. 2003;137(4):617-20.

Research Article

¹McGill University, Montreal, QC, Canada²Lamont-Doherty Earth Observatory, Columbia University, New York, USA

Keywords

Multi-year ice (MYI), First-year ice (FYI), Promotion, Demotion, Export

Email Correspondence

jamie.hart@mail.mcgill.ca

Jamie Hart¹, Bruno Tremblay^{1,2}, Charles Brunette¹, Carolina Dufour¹, Robert Newton²

Examining the Transition from a perennial to a seasonal sea ice cover in the Arctic Ocean: A Lagrangian Approach

Abstract

Background: Declining Arctic sea ice extent has been accompanied by a large loss in multiyear ice (MYI). The dynamic and thermodynamic processes which affect this transition include promotion of first year ice (FYI) to MYI, demotion (melting) of MYI to open water, and ice export through Fram Strait. In this study we quantify the relative importance of these three processes.

Methods: We use the Lagrangian Ice Tracking System which employs satellite-derived sea ice drift vectors combined with sea ice concentrations to find annual areas of promotion, demotion, and export.

Results: Over the satellite record (1989-2015), we quantify the total contributions to sea ice extent loss from promotion (+30 million km²), demotion (-19.7 million km²), and export of MYI (-18.6 million km²). The result is a total net loss of 8.3 million km² of MYI. We find that all three processes are positively correlated with minimum sea ice extent and are increasing with rates of +0.165 million km²/decade, -0.146 million km²/decade, and -0.096 million km²/decade for promotion, demotion, and export respectively. We also compute the negative ice growth feedback at 0.59 (with $r^2=0.27$). This indicates that ice pack recovers, on average, 59% of the MYI area lost to demotion/export through promotion of FYI the following winter.

Limitations: Uncertainties in the drift speed are compounded by the weekly temporal resolution of the model, which affects the resulting estimates of demotion and promotion area.

Conclusion: Demotion and export combined are increasing faster than promotion and represent a larger area contribution. This imbalance accounts for the observed loss of MYI area.

Introduction

Arctic sea ice extent has been declining rapidly over the satellite record. From 1979 to 2014, the September minimum sea ice extent has declined at a rate of -13.3% per decade (1). The reduced spatial coverage has been accompanied by a decrease in multi-year ice (MYI; 2,3,4,5), defined as ice that has survived through at least one melt season. In the mid-1980s, MYI constituted 75% of the total ice extent (3) but has fallen by more than 50% since the year 2000 (5). As first year ice (FYI) replaces MYI as the dominant ice type, the ice pack is generally thinner and shows reduced extent for a given summer melt (5).

Promotion of FYI to MYI and demotion of MYI to open water have critical implications for marine ecosystems. Plankton, nutrients and sediment are entrained into FYI as it forms and transported to the site where the ice melts. Promotion to MYI is important for the long-distance transport of these materials, e.g.: to the Central Arctic and other peripheral seas. Indeed, this transport results in the high productivity of some Arctic peripheral sea fisheries, such as the Barents Sea (6). Recently, changes in FYI and MYI distribution are altering this transport and endangering the ecosystem. This ecosystem is estimated to disappear within 20 years, the most rapid disappearance of an ecosystem associated with climate change yet recorded (7).

Demotion is not the only mechanism of MYI loss: export through the Fram Strait and subsequent melting also diminishes Arctic MYI area. The Fram Strait is the passage between Greenland and Svalbard. Fram Strait ice export is important for the Arctic sea ice budget, since about 10% of the total sea ice mass is exported through here annually (8). Ice export and import through other straits (Bering, Davis, Nares) are an order of magnitude smaller (9,10,11). Moreover, export rates through the Fram Strait are determined by the strength and position of the Transpolar Drift Stream (TDS) and can vary annually by up to 50% (12). On longer times-

cales, Fram Strait export is affected by atmospheric climate variations such as the North Atlantic Oscillation (NAO)/Arctic Oscillation (AO). During a positive NAO/AO, the pressure gradient across the Fram Strait increases, causing higher export (13).

When the Arctic system is perturbed by external climate forcings, various internal processes will react either by pushing the system back to equilibrium (negative feedback) or further away from its equilibrium (positive feedback). There are several positive feedbacks which amplify warming once open water and FYI are present. For example, the ice-albedo positive feedback, wherein warming decreases ice cover or increases meltpond fraction, creating a surface with a lower albedo (reflectivity), which absorbs more solar energy, leading to more warming (14). Nevertheless, there are also negative feedbacks. For example, the ice-growth feedback, where open water or thinner ice cover both emit more radiation than colder thicker ice (thin ice conducts the ocean's heat), causing more heat loss, and therefore ice growth in the next year. Overall, the Arctic's positive feedback loops dominate and drive the transition from a large area of thick MYI towards thinner FYI and a seasonally ice-free Arctic (15,16).

The occurrence of an ice-free arctic summer is currently predicted to occur before the end of the 21st century (4). In 2007, results from the Coupled Model Intercomparison Project Phase 3 (CMIP3) reported in the United Nations Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) predicted nearly-ice free conditions in the year 2075. More recent estimates now predict occurrence of an ice-free Arctic in just a few decades, with many CMIP5 models showing essentially ice-free conditions by 2050 (17). There are still large uncertainties in these estimates due to uncertainties in the physics of climate models, future anthropogenic emissions, and natural climate variability. Indeed, natural climate variability alone creates uncertainties on the order of two decades (18). Therefore, understanding the processes of MYI loss is imperative to improving climate models' performance and decreasing uncertainty in the prediction of an ice-free Arctic summer.

McGill Science Undergraduate Research Journal - msurj.com

In this study, we investigate the large loss of MYI in the Arctic Ocean, with the goal of quantifying the relative importance of the thermodynamic and dynamic processes causing this change, namely: demotion (melting) of MYI to open water, promotion of FYI to MYI, and export of MYI through the Fram Strait. We also map the areas of the Arctic where these processes predominantly occur. To accomplish this, we use the Lagrangian Ice Tracking System (19-20) forced with satellite-derived sea ice drift data together with observed sea ice concentration. The results from this study will help determine if the mechanisms responsible for the transition to a seasonal ice cover in current general circulation models (GCMs) are realistic. If not, it will help identify key processes that are not currently well-represented in GCMs and limit seasonal sea ice predictability.

Methods

I. Sea Ice Concentration

Sea ice concentration (SIC) is defined as the fraction of each grid cell covered by ice. Furthermore, the sea ice extent is the total area covered by the sea ice pack, with each grid cell requiring at least 15% SIC to be accepted as ice-covered. We justify this commonly-used threshold by the fact that concentration increases quickly from the edge of the pack ice inwards. We calculate the sea ice extent from the SIC data and the grid cell area. The SIC data comes from the National Oceanic and Atmospheric Administration (NOAA)/National Snow & Ice Data Center (NSIDC) Climate Data Record (CDR) Passive Microwave Sea Ice Concentration Version 3 (21), which provides daily and monthly SIC data. This information is provided on a polar stereographic grid with a 25 x 25 km Equal-Area Scalable Earth (EASE) grid spatial resolution (22). We produce weekly SIC averages interpolated on the EASE-Grid from 1980 to 2015. Note that in 1987 and 1988, due to maintenance issues, two periods of data were lost: 1987 weeks 28-31, and 1987 weeks 49-52 through 1988 weeks 1-2. The second period through the 87/88 winter could be set to the surrounding week's SIC, since SIC changes little over the winter in the central Arctic. However, the first period in the spring of 1987 is a time of large variability in sea ice: therefore, we simply exclude 1987 in our results.

II. Sea Ice Drift

Sea ice drift is the motion of sea ice due to wind, ocean currents, the Coriolis force, internal ice stress, and sea surface tilt. Only sea ice which is attached to the shore ("fast ice") does not drift. We use the NSIDC Polar Pathfinder (PPF) Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 3 (23), which provides daily, monthly, and weekly sea ice drift data on a 25 x 25 km EASE grid projection. The sea ice drift vectors are the product of an interpolation which combines input from four types of sources: buoys from the International Arctic Buoy Program (IABP), passive microwave sensors (AMSR-E, SMMR, SSM/I, SSMIS), visible and infrared radiometer channels (AVHRR), and free-drift estimate derived from NCEP/NCAR wind Reanalysis (23). When the aforementioned data is missing or unreliable, we use free drift estimates derived from NCEP/NCAR Reanalysis data. We use the weekly averaged sea ice drift vectors from 1980 to 2015.

III. Lagrangian Ice Tracking System (LITS)

Sea ice was advected (i.e. transported horizontally) using the Lagrangian Ice Tracking System (LITS) (19-20). This system takes an initial ice-covered tracer on the EASE grid and advects it in one-week increments using interpolated PPF weekly drift vectors from the start date to the end date. Through each advection step, LITS checks the SIC to make sure the tracer is still in an area covered by sea ice (we refer to this tracer as "active"). If the tracer melts (SIC drops below 15%), then LITS marks it as inactive and stops advection. The location error (i.e. the distance between the true buoy location and the trajectory estimated by LITS) yields a median and third quartile error of 7% and 16%, respectively, for typical Arctic Ocean sea ice drift speeds of 3-5 cm/s (20).

IV. Advection Protocol

We show a step-by-step example for the advection procedure from year 2014 to 2015 (Fig. 1). We start each year at the week of the minimum

sea ice extent (SIE), which occurs at the end of the summer melt season, usually in September. According to our definition, all sea ice present at this time is MYI since it has survived the summer melt. The initial 2014 minimum sea ice pack is tagged inside of a mask which excludes the Canadian Arctic Archipelago (CAA, where no winter drifts are present) and everything south of the Fram Strait (Fig. 1a), with each tracer representing a 25 x 25 km grid cell. The boundary is then found using the MATLAB function *bwboundaries*, taking only the largest object found and excluding interior holes. The tracers are then advected in weekly increments until the time of the 2015 minimum sea ice extent (Fig. 1b). If a tracer melted early in the season (i.e. after March, at the very end of winter), its last position is recorded. This allows us to identify ice loss associated with the melt front early in the spring/summer season. Because of the dispersion of the tracers, the new boundary is less well-defined. In this case, we use the MATLAB function *alphaShape* to bound the advected area (Fig. 1c). The *alphaShape* boundary is considered the 2014 ice edge advected to September 2015 (Fig. 1d). Next, the 2015 observed minimum sea ice edge is found using *bwboundaries* as before and is overlaid on top of the 2014 advected edge (Fig. 1d). Finally, by subtracting the two boundaries, the areas of promotion and demotion are identified (Fig. 1e), indicating regions of FYI promotion and MYI demotion. Where the observed ice edge extends further south than the previous year advected ice edge, newly formed ice survived the 2015 summer melt: this ice corresponds to the promotion of FYI to MYI. Where the observed ice edge extends farther north than the advected ice edge, the existing MYI is lost through melting: this ice corresponds to the demotion of MYI to open water. Note that FYI promoted to MYI could be demoted (melt or export) on its second year of life.

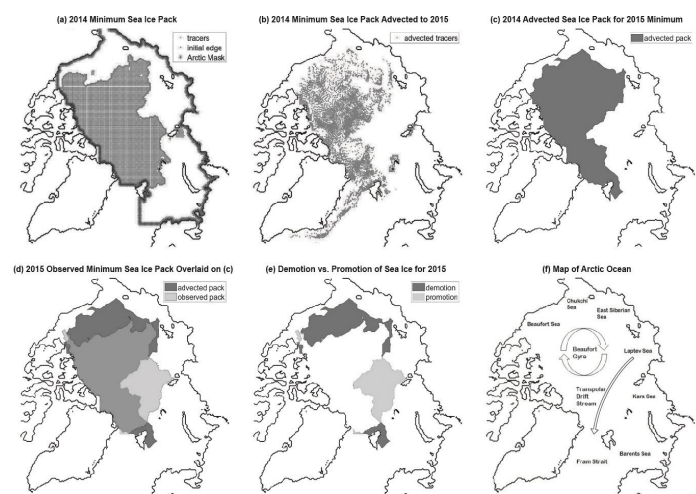


Fig. 1. (a) Initial position of tracers for 2014 pack. (b) Final position of tracers in (a) advected 1 year to September 2015. (c) Advected MYI pack from 2014 to 2015. (d) 2015 minimum pack overlaid on (c), where medium grey tone indicates region overlap. (e) Regions of promotion and demotion. (f) Arctic map listing peripheral seas/sea ice circulation features.

V. Fram Strait MYI Export

MYI is also lost by leaving the Arctic region (mask) through the Fram Strait, which is not properly accounted for by the method described above. We know that once this MYI crosses the Fram Strait, it will eventually be melted by the North Atlantic heat flux. Therefore, we count it as a loss outside the region. Hence, we need to calculate the transport of MYI across Fram Strait accurately. When the method described in section 2. IV is applied to the Fram Strait outside the mask, MYI loss is underestimated by almost 80% compared with other studies (8,12). We explain this by the North Atlantic heat influx continually melting parcels in the same location, and this overlapping of melted parcels resulting in a smaller computed area boundary compared to the actual area of MYI exported. As such, we require a different approach when addressing Fram Strait export. Instead, we count the number of tracers advected out the Fram Strait each year and multiply by the area of a grid cell (625 km²). Since each of these tracers came from the Arctic Ocean and occupied an area of 25 x 25 km

begin with, we have a direct way to calculate the loss of Arctic Ocean MYI area. Using our method, we find a MYI export over the 35-year period of 18.6 M km^2 , or $547\,000 \text{ km}^2/\text{year}$. Our result is approximately 60% of the export computed by other studies who considered FYI and MYI (8,12). This is appropriate, since over the last 8 years the proportion of Fram Strait export that comes from MYI ranges from 64% to 90% of the total export, where the remaining percentage consists of FYI (24).

VI. Error analysis

We estimate LITS ice-drift track errors by comparing LITS, forced with PPF sea-ice drifts with buoy drift trajectories, which we consider to be error-free. The result is a median location error of 7% for typical Arctic sea ice drifts of 3-5 cm/sec (20). There is also an error in the areas of FYI promotion and MYI demotion associated with the temporal resolution of the model (one week). An upper bound error estimate for the error of tracers starting near the ice edge and drifting perpendicular to the ice edge towards open water can be written as $u_{ice} * \Delta t * L$ where Δt is 7 days and L is a typical synoptic length scale (500 km). Such a tracer would end its trajectory in open water and would be tagged as inactive for the next time step. As such, this error only applies to the demotion of MYI and always leads to an overestimate in MYI demotion area. For instance, an average ice drift speed of 4 or 10 cm/s in the central Arctic yields an upper bound error of approximately 12 or 30 thousand km^2 .

Results and Discussion

I. Regions of Promotion and Demotion

We find two unique patterns for both promotion and demotion in the periods 1980-1999 and 2000-2015 (Fig. 2). We compare observations from the beginning of the satellite era (1980-1999) to the more recent period (2000-2015). In the pre-2000 (1980-1999), the regions of promotion and demotion are spread among the peripheral seas (Fig. 2a and 2b), whereas in the post-2000 (2000-2015), these regions are larger and extend poleward because of the increased drift speed (25, 26, Fig. 2c and 2d).

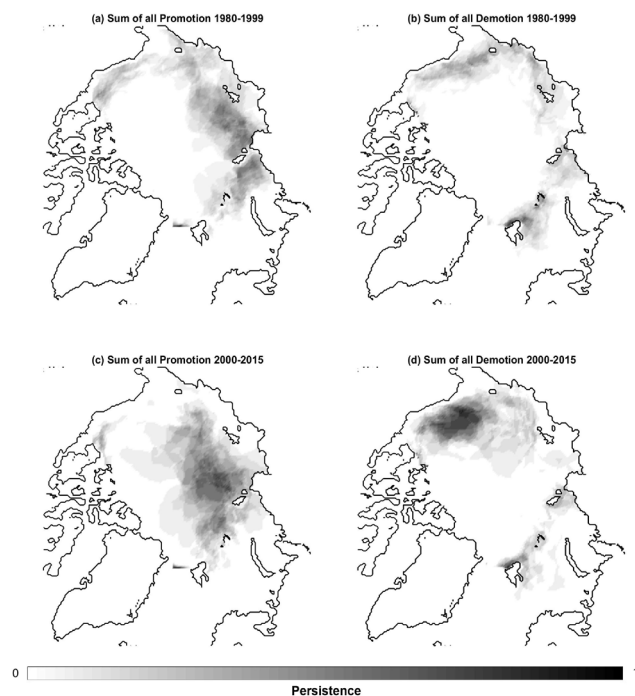


Fig. 2. Regions of (a,c) promotion and (b,d) demotion for (a,b) 1980-1999 and (d,c) 2000-2015. The grey gradient indicates persistence, which refers to the fraction of the given time period that a pixel experienced (a,c) FYI promotion or (b,d) MYI demotion.

In the earlier period, promotion is mainly located in the Laptev Sea extending into the Kara Sea, with smaller centres also along the shorelines in

the East Siberian, Chukchi, Beaufort, and Northern Barents Seas (Fig. 2a). It is well documented that promotion occurs consistently in the Laptev Sea due to divergence from the coastline, giving it the name of sea ice factory (27). This region is the source of the Transpolar Drift Stream, which transports FYI from the Laptev sea to the central Arctic and Fram Strait (20). In the later period, this main promotion region increases in size and persistence (Fig. 2c), indicating a stronger transpolar drift stream (25,26).

In the earlier period, promotion is mainly located in the Laptev Sea extending into the Kara Sea, with smaller centres also along the shorelines in the East Siberian, Chukchi, Beaufort, and Northern Barents Seas (Fig. 2a). It is well documented that promotion occurs consistently in the Laptev Sea due to divergence from the coastline, giving it the name of sea ice factory (27). This region is the source of the Transpolar Drift Stream, which transports FYI from the Laptev sea to the central Arctic and Fram Strait (20). In the later period, this main promotion region increases in size and persistence (Fig. 2c), indicating a stronger transpolar drift stream (25,26).

In general, the transition from a perennial to a seasonal ice cover leads to a clear separation between centers of demotion and promotion. More specifically, promotion mainly occurs in the Eurasian Arctic, while demotion (of primarily Beaufort Sea MYI) is more prominent in the Pacific sector of the Arctic (Fig 2c and 2d). Although areas of promotion and demotion have become more segregated over time, we still find overlap in peripheral regions, most notably in the Central Arctic Ocean and East Siberian Sea. We also identify small areas of intersection along the Beaufort Sea, Laptev Sea, and around the islands in the Kara and Barents Seas.

II. Temporal Variability

We observe an approximate balance between promotion, demotion and export in the 80s and 90s, but an increasing trend in promotion as well as demotion and export since 2000 (Fig. 3). This pattern shift could be related to atmospheric changes in NAO/AO, a serious consideration which is left to future work. Over the 35-year period, the sum of all promotion is 30.0 million km^2 , whereas the sum of all demotion is -19.7 million km^2 , and the sum of all export is -18.6 million km^2 . The net change in MYI area (i.e. promotion minus demotion minus export) for the entire period is -8.3 million km^2 , which corresponds to $-2.4 \text{ million km}^2/\text{decade}$. In the later period (2000-2015) we calculate MYI loss at $-5.5 \text{ million km}^2$, which is much higher than satellite-derived estimates of -2 million km^2 for 1999-2017 (5). Our very different approaches, lagrangian tracking and eulerian satellite field analysis, may explain this discrepancy. We find increasing trends in promotion ($0.16 \text{ million km}^2/\text{decade}$), demotion ($-0.15 \text{ million km}^2/\text{decade}$), and export ($-0.096 \text{ million km}^2/\text{decade}$), all beyond the 98% confidence level. This results in a trend for the net change in MYI of $-0.08 \text{ million km}^2/\text{decade}$, with a significance of 72%.

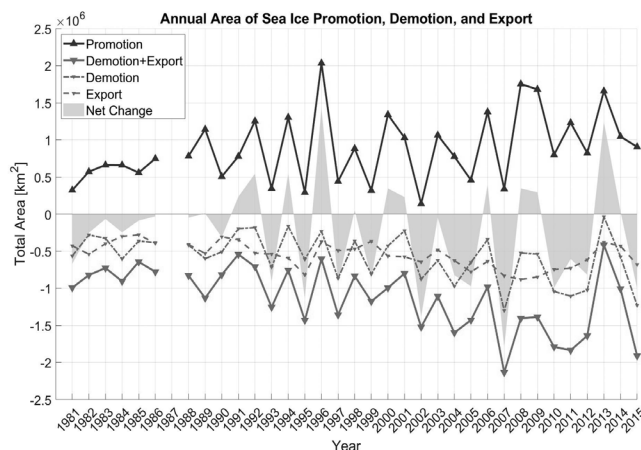


Fig. 3. Time series of sea ice area of FYI promotion (upward arrow), export (long dashed gray), MYI demotion (short dashed gray), and MYI demotion plus export (downward arrow) for time period 1981-2015. The lighter filled area represents the net change in sea ice area from year $n-1$ to year n (promotion - demotion - export).

III. Correlations

We further investigate the temporal variability of promotion, demotion, and export by examining the correlation of SIE minimum. We compute the annual correlation of detrended anomaly time series for promotion, demotion, and export against the same year minimum SIE. We find moderate positive correlations of 0.52, 0.51 and 0.30, respectively. Therefore, promotion and demotion serve as diagnostic variables for increasing sea ice loss. We attribute decreasing minimum sea ice extent to increasing rate of demotion and export. This will be further investigated.

IV. Negative Ice Growth Feedback

We analyze both positive and negative feedbacks using lagged cross-correlations between demotion, promotion, and export. The only significant correlation we find is the negative ice growth feedback, which is the tendency for a year of high demotion and export to be followed by a year of high promotion. To quantify the negative ice growth feedback, we investigate the relationship between MYI loss mechanisms (demotion and export) and the subsequent year's MYI growth mechanism (promotion). The best line fit between scatter plot of promotion in year $n+1$ versus demotion and export in year n has a slope of 0.59 with an r -squared value of 0.27, indicating a moderate strength feedback (Fig. 4). This means that on average 59% of MYI demotion in a given year will be compensated by FYI promotion the following year.

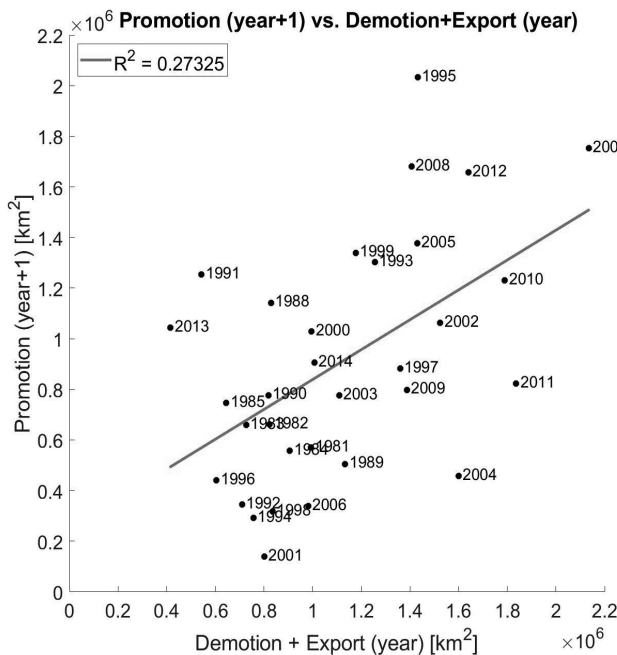


Fig. 4. Scatter plot of the total sea ice area of promotion in year $n+1$ versus demotion + export for year n . The gray dotted line is a linear regression with an r^2 value of 27%. Years 1986, 1987, and 2015 cannot be calculated because of the missing 1987 data (see section 2.1) and the end of the dataset in 2015.

Given the feedback loops present in the perturbed Arctic system, we cannot assume any current trend to continue linearly. If the increase in FYI and decrease in MYI continue, the mechanisms of warming and ice loss may intensify and increase the rate exponentially.

Conclusion

In this study, we examine the mechanisms responsible for recent MYI loss in the Arctic. To do so, we use a Lagrangian approach to quantify promotion of FYI to MYI, demotion of MYI to open water, and export of MYI through the Fram Strait. Results show a change in both the MYI growth and loss in the Arctic. Since 1980, we estimate promotion of FYI to MYI (8.6 million $\text{km}^2/\text{decade}$ for a total of 30 million km^2), demotion of MYI to open water (-5.6 million $\text{km}^2/\text{decade}$ for a total of -19.7 million km^2),

and export of MYI through the Fram Strait (-5.3 million $\text{km}^2/\text{decade}$ for a total of -18.6 million km^2). The combined effect of demotion and export is a net MYI loss of -8.3 million km^2 over the period. This is in general agreement with satellite derived estimate of sea ice extent in the Arctic since the beginning of the satellite era in the late seventies. We find increasing trends for each process at $+0.165$ million $\text{km}^2/\text{decade}$ for promotion, -0.146 million $\text{km}^2/\text{decade}$ for demotion, and -0.096 million $\text{km}^2/\text{decade}$ for export. We document a correlation between high demotion and export years and subsequent high promotion years, corresponding to a negative ice growth feedback of 59%. We find that demotion and promotion are both positively correlated with minimum sea ice extent. We thus conclude that demotion is primarily responsible for the transition from perennial to seasonal ice pack. Future work will include considering potential biases in sea ice drift vectors and investigating to what extent they can cause under or over-estimation of demotion and promotion areas.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council's Undergraduate Student Research Award. We thank NOAA and NSIDC for the SIC data Climate Data Record (CDR), as well as NSIDC for the Polar Pathfinder sea ice drift data.

References

1. Serreze MC, Stroeve J. Arctic sea ice trends, variability and implications for seasonal ice forecasting. *Phil. Trans. R. Soc. A.* 2015 Jul 13;373(2045):20140159.
2. Comiso JC. A rapidly declining perennial sea ice cover in the Arctic. *Geophysical Research Letters.* 2002 Oct 18;29(20):17-1-17-4.
3. Maslanik J, Stroeve J, Fowler C, Emery W. Distribution and trends in Arctic sea ice age through spring 2011. *Geophysical Research Letters.* 2011 Jul 14;38(13).
4. Stroeve JC, Kattsov V, Barrett A, Serreze M, Pavlova T, Holland M, et al. Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophysical Research Letters.* 2012 Aug 28;39(16).
5. Kwok R. Arctic sea ice thickness, volume, and multiyear ice coverage: losses and coupled variability (1958–2018). *Environmental Research Letters.* 2018 Oct 12;13(10):105005.
6. Barber DG, Hop H, Mundy CJ, Else B, Dmitrenko IA, Tremblay JE, et al. Selected physical, biological and biogeochemical implications of a rapidly changing Arctic Marginal Ice Zone. *Progress in Oceanography.* 2015 Dec 1;139:122-50.
7. Fossheim M, Primicerio R, Johannesen E, Ingvaldsen RB, Aschan MM, Dolgov AV. Recent warming leads to a rapid borealization of fish communities in the Arctic. *Nature Climate Change.* 2015 May 18;5(7):673-7.
8. Smedsrud LH, Halvorsen MH, Stroeve JC, Zhang R, Kloster K. Fram Strait sea ice export variability and September Arctic sea ice extent over the last 80 years. *The Cryosphere.* 2017 Jan 13;11(1):65-79.
9. Kwok R. Baffin Bay ice drift and export: 2002–2007. *Geophysical Research Letters.* 2007 Oct 2;34(19).
10. Kwok R, Cunningham GF, Wensnahan M, Rigor I, Zwally HJ, Yi D. Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008. *Journal of Geophysical Research: Oceans.* 2009 Jul 7;114(C7).
11. Babb DG, Galley RJ, Asplin MG, Lukovich JV, Barber DG. Multiyear sea ice export through the Bering Strait during winter 2011–2012. *Journal of Geophysical Research: Oceans.* 2013 Oct 1;118(10):5489-503.
12. Kwok R, Cunningham GF, Pang SS. Fram Strait sea ice outflow. *Journal of Geophysical Research: Oceans.* 2004 Jan 6;109(C1).

13. Kwok R, Rothrock DA. Variability of Fram Strait ice flux and North Atlantic oscillation. *Journal of Geophysical Research: Oceans*. 1999 Mar 15;104(C3):5177-89.
14. Curry JA, Schramm JL, Ebert EE. Sea ice-albedo climate feedback mechanism. *Journal of Climate*. 1995 Feb 1;8(2):240-7.
15. Auclair G, Tremblay LB. The Role of Ocean Heat Transport in Rapid Sea Ice Declines in the Community Earth System Model Large Ensemble. *Journal of Geophysical Research: Oceans*. 2018 Nov 29;123(12):8941-57.
16. Holland MM, Bitz CM, Tremblay B. Future abrupt reductions in the summer Arctic sea ice. *Geophysical research letters*. 2006 Dec 12;33(23).
17. Overland JE, Wang M. When will the summer Arctic be nearly sea ice free?. *Geophysical Research Letters*. 2013 May 28;40(10):2097-101.
18. Jahn A, Kay JE, Holland MM, Hall DM. How predictable is the timing of a summer ice-free Arctic?. *Geophysical Research Letters*. 2016 Sep 16;43(17):9113-20.
19. Williams J, Tremblay B, Newton R, Allard R. Dynamic preconditioning of the minimum September sea-ice extent. *Journal of Climate*. 2016 Aug 3;29(16):5879-91.
20. DeRepentigny P, Tremblay LB, Newton R, Pfirman S. Patterns of sea ice retreat in the transition to a seasonally ice-free Arctic. *Journal of Climate*. 2016 Sep 29;29(19):6993-7008.
21. Meier W, Fetterer F, Savoie MH, Mallory S, Duerr R, Stroeve JC. NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 3. NSIDC: National Snow and Ice Data Center: Boulder, Colorado USA 2017 [cited 12 Feb 2019]. Available from: <https://nsidc.org/data/nsidc-0116>
22. Peng G, Meier WN, Scott DJ, Savoie MH. A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring. *Earth System Science Data*. 2013 Oct 15;5(2):311-8.
23. Tschudi M, Fowler C, Maslanik J, Stewart S, Meier W. Polar Pathfinder Daily 25 km EASE-Grid Sea Ice Motion Vectors, Version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center: Boulder, Colorado USA 2016 [cited 12 Feb 2019]. Available from: <https://nsidc.org/data/G02202>
24. Ricker R, Girard-Ardhuin F, Krumpen T, Lique C. Satellite-derived sea ice export and its impact on Arctic ice mass balance. *The Cryosphere*. 2018 Sep 25;12(9):3017-32.
25. Kwok R, Spreen G, Pang S. Arctic sea ice circulation and drift speed: Decadal trends and ocean currents. *Journal of Geophysical Research: Oceans*. 2013 Apr 9;118(5):2408-25.
26. Rampal P, Weiss J, Marsan D. Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979–2007. *Journal of Geophysical Research: Oceans*. 2009 May 14;114(C5).
27. Reimnitz E, Dethleff D, Nürnberg D. Contrasts in Arctic shelf sea-ice regimes and some implications: Beaufort Sea versus Laptev Sea. *Marine Geology*. 1994 Jul 1;119(3-4):215-25.
28. Onarheim IH, Eldevik T, Smedsrud LH, Stroeve JC. Seasonal and Regional Manifestation of Arctic Sea Ice Loss. *Journal of Climate*. 2018 Jun 15;31(12):4917-32.
29. Kwok R, Cunningham GF. Contribution of melt in the Beaufort Sea to the decline in Arctic multiyear sea ice coverage: 1993–2009. *Geophysical Research Letters*. 2010 Oct 1;37(20).

Research Article

¹Fisheries & Aquatic Sciences Program, School of Forest Resources & Conservation, University of Florida

²Department of Geography, University of Florida

³Geomatics, School of Forest Resources & Conservation, University of Florida

Keywords

Giant Kelp, Geomorphometry, Habitat Mapping, Invasive Species, Species Distribution Modeling

Email Correspondence

vlecours@ufl.edu

Michael Espriella¹, Tyler Schaper², Alison Atchia¹, Katherine Rose¹, Vincent Lecours^{1,3}

Habitat mapping of giant kelp (*Macrocystis pyrifera*) and devil weed (*Sargassum horneri*) off the coast of Santa Catalina Island, California

Abstract

Background: *Macrocystis pyrifera*, commonly known as giant kelp, is a fast-growing brown alga that typically inhabits temperate waters. In southern California, *M. pyrifera* provides many ecologically and economically significant ecosystem services. *Sargassum horneri*, a non-native brown macroalga commonly known as devil weed, often outcompetes *M. pyrifera* while providing fewer ecological or economical benefits. Examining potential areas of species overlap is key to understanding the invasion potential of *S. horneri* and essential to the implementation of removal efforts. This study aims to map the suitable habitat of *M. pyrifera* and invasive *S. horneri* in the coastal waters of Santa Catalina Island, California, and to quantify any overlapping habitat between the two macroalgae.

Methods: Broadly defined potential habitats were characterized around Santa Catalina Island using an unsupervised approach to habitat mapping based on a series of abiotic surrogates mapped at a 2 m spatial resolution. *In situ* substrate data were then overlaid onto the unsupervised classification to identify spatial associations between substrate type and potential habitats, and to interpret the classes. To predict the distribution of *M. pyrifera* and *S. horneri* around Santa Catalina Island based on their respective association with the environment, maximum entropy (MaxEnt) was used to produce species distribution models. The resulting models for *M. pyrifera* and *S. horneri* were overlaid to identify potential areas of conflict based on suitable habitat overlap.

Results: The unsupervised approach to habitat mapping resulted in a map of four potential habitats around Santa Catalina Island based on substrate cover. Sand was the most dominant type of substrate. The supervised approach using MaxEnt identified 10.27% of the study area as suitable habitat for *M. pyrifera* and 7.37% as suitable habitat for *S. horneri*. A total of 33.56% of the suitable habitat for *M. pyrifera* was found to also be suitable for *S. horneri*.

Limitations: The characterization of habitats and the species distribution modeling were limited to the study of benthic terrain characteristics due to the unavailability of other high-resolution environmental data (e.g., hydrodynamics and chemical data) around Santa Catalina Island. In addition, data were not available for the very shallow waters near the coast, where giant kelp is often found. Given the complexity of this ecosystem, the addition of other variables and data coverage closer to the coast would potentially make the maps and models more representative of the actual distribution of *M. pyrifera* and *S. horneri* and provide a more complete understanding of their environmental preferences.

Conclusion: This study provides insight into the kelp forest ecosystems found in California's Channel Islands; it is a vital first step in order to understand the potential areas for invasion of *M. pyrifera* by *S. horneri*, thus supporting decision making and efforts to control *S. horneri* abundance.

Introduction

Macrocystis pyrifera, commonly known as giant kelp, is a fast-growing brown alga normally found in temperate, nutrient-rich waters off the west coast of the Americas and the southern coast of Australia. It is most well-known for its massive size and plant-like appearance. Giant kelp begins its growth cycle by anchoring itself to a suitable substrate with a root-like holdfast that does not absorb nutrients from the benthic environment (1). From there it grows towards the ocean surface, reaching anywhere from 14 to 40 m using gas-filled sacs – the pneumatocysts – along the stipe to remain upright in the water column (2). Blades sprouting from the pneumatocysts maximize the area available for photosynthesis. The combination of this complex structure and the rapid growth rate of *M. pyrifera* makes it a foundational species in its ecosystem. For instance, it provides a three-dimensional habitat for other species, and the high rate of photosynthesis that takes place along the thallus incorporates large amounts of inorganic carbon into the food-web (3). In southern California, *M. pyrifera* supports a complex food-web consisting of nearly 300 species, many of which are ecologically and economically significant (4). Some significant markets include the recreational rockfish, the kelp bass *Paralabrax clathratus*, and the

commercial lingcod fisheries (4).

Santa Catalina Island, which is located about 40 km from the Los Angeles shoreline in the Pacific Ocean (Fig.1), is one of California's most well-known rocky Channel Islands. Although some development has allowed Santa Catalina Island to capitalize on fisheries and tourism, a large majority of the island remains uninhabited, with a large population of its residents living within a small area. Santa Catalina Island possesses several unique ecosystems (e.g., mountainous habitats, grasslands, oak woodlands), including a variety of coastal ecosystems such as sand dunes, beaches, and giant kelp forests. Many of the 900,000+ annual visitors to Santa Catalina Island – compared to the 4,000 residents of the island – come to dive, fish or take a boat tour in the kelp forests (5). These activities, which translate into \$167 million in annual spending and over 1,200 jobs on the island (5), put disproportionate pressures on the kelp ecosystem and its species: in the waters surrounding Santa Catalina Island, more than 150 recorded fish species spend a portion of their life cycle in kelp forests using it either as a food source or habitat (4).

A suite of stressors, namely sea star wasting disease, warmer waters, and

rising sea urchin populations have recently threatened southern California's giant kelp population (6). Aerial surveys demonstrate that the combination of these factors has led to a decline in the distribution of *M. pyrifera*, leading to a loss of carbon fixation and suitable habitat for other species (7). On Santa Catalina Island, these stressors are exacerbated by the non-native *Sargassum horneri*, which often outcompetes *M. pyrifera*. Similar to *M. pyrifera*, *S. horneri* is a brown macroalga that establishes a

Catalina Island includes nine marine conservation areas (11), spatially-explicit approaches to *S. horneri* removal based on distribution and invasion patterns have not been included in management practices (12). The production of both giant kelp and devil weed habitat suitability maps are needed to identify specific areas to focus devil weed removal efforts or isolate and protect giant kelp. Spatial analysis of *M. pyrifera* and *S. horneri* can provide preliminary information on the likelihood of invasion in vulnerable areas by identifying overlap in suitable habitats for both species. The information gained about the habitat distribution of each species could be used to develop a comprehensive response effort. This study aims to map the suitable habitats of *M. pyrifera* and *S. horneri* in the coastal waters of Santa Catalina Island and to quantify overlapping suitable habitats in order to provide a decision-making tool to aid efforts and resource allocation in the control of *S. horneri* abundance.

Methods

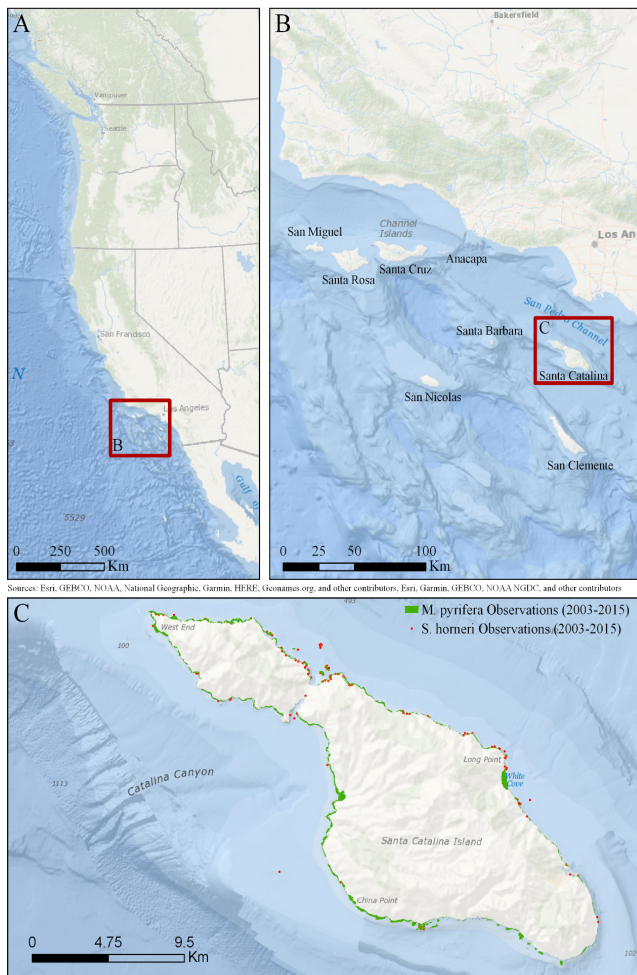


Figure 1. West Coast of the United States of America (A), with the location of the eight Channel Islands (B) and Santa Catalina Island with *M. pyrifera* and *S. horneri* observations from 2003 to 2015 (C)

holdfast on the ocean floor and uses pneumatocysts to remain upright in the water column. *Sargassum horneri* is native to temperate regions off of Korea and Japan. It was first discovered in the Baja region of California in 2003 and in waters surrounding Santa Catalina Island in 2006. A common hypothesis used to explain the macroalga's arrival is that it was transported by commercial vessels (e.g., in ballast water) (8). Its rapid expansion from the South to the North is partly attributed to its prolific reproductive strategy; *S. horneri* generates and sheds hundreds of germinating embryos that can settle or survive the journey to neighboring ecosystems (8).

In an already stressful environment for *M. pyrifera*, the faster growth rate of *S. horneri* inhibits the recovery of *M. pyrifera* population to its historical abundance and distribution by outcompeting it for space and light (9). Within the first year of its discovery, *S. horneri* has spread along the leeward side of Santa Catalina Island before progressing to the windward side of the island in 2007, consistently forming dense groves of roughly 100 adult individuals and 1,000 immature individuals per square meter (8,10). *Sargassum horneri* cannot provide the same habitat as the tall, canopy-forming *M. pyrifera*, and thus the recruitment of fish species such as kelp bass is negatively impacted by the non-native *S. horneri* invasion (9). While current management of the marine habitats surrounding Santa

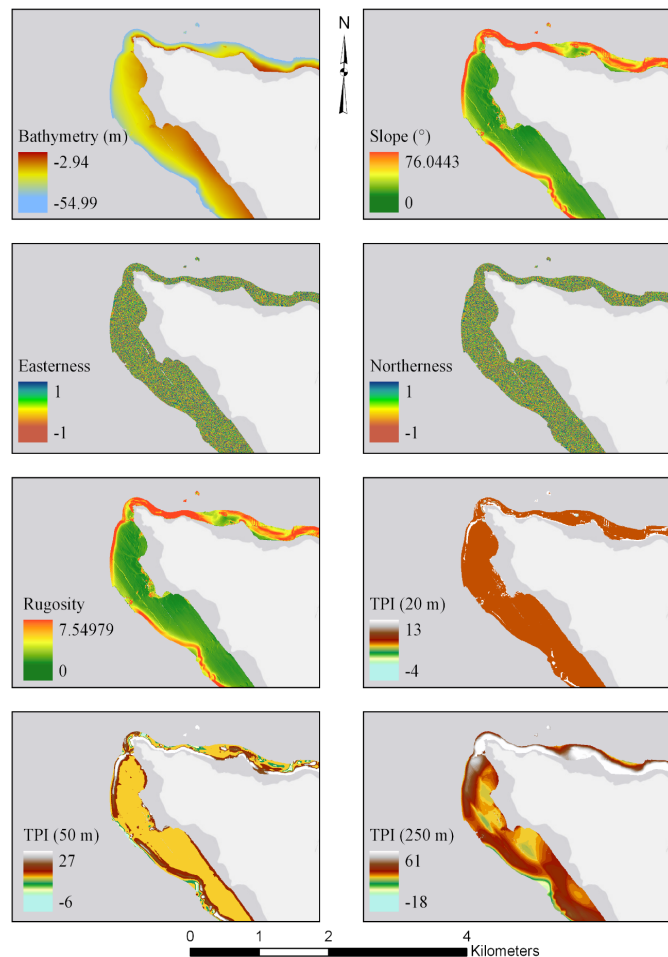


Figure 2. Example of each variable used in this study, from the northwestern tip of Santa Catalina Island. TPI stands for topographic position index.

Abiotic data of the seafloor were accessed through the Seafloor Mapping Lab at California State University Monterey Bay (13), and included bathymetry (i.e., depth of the seafloor), which was collected in 2008 using acoustic remote sensing technologies. Bathymetry derivatives, namely slope, aspect, topographic position indices (TPI), and rugosity, were also provided by the Seafloor Mapping Lab at California State University Monterey Bay (13). Aspect, which represents the compass orientation of the slope and can be a proxy for hydrodynamics, was converted to northerness (cosine of aspect) and easternness (sine of aspect) to remove circularity in the data and make them statistically valid (14). TPIs, which indicate whether a given location is higher or lower than its immediate surroundings, was available at three different spatial scales, 20 m, 50 m, and 250 m, thus potentially capturing seafloor features at different scales and of

different sizes. Because the deepest occurrence of *M. pyrifera* was found at a depth of 55 m (15), we masked all the data corresponding to waters deeper than that threshold. We note, however, that most often, giant kelp is found in waters between 10 and 40 m, depending on water clarity and light penetration (15). The total area covered by the data was about 35.5 km². An example of each variable is shown in Fig. 2. Correlations between the data layers were quantified to inform data selection.

We used two different strategies for habitat mapping according to Brown *et al.* (16) to model suitable habitats for giant kelp and devil weed. First, to characterize broad potential habitats around Santa Catalina Island based on abiotic environmental characteristics, we used an unsupervised approach to habitat mapping (16). The selected data were inputted into a modified k-means classification tool in the Whitebox GAT v. 3.4.0 software (17). K-means is the most commonly used clustering algorithm that partitions datasets into meaningful classes (18). The maximum number of iterations and the cluster merger distance were set to 10, the pixel class change threshold was set to 10%, and the minimum allowable class size was set to 10 pixels. The cluster centers were initialized with maximum dispersion along the diagonal. The resulting unsupervised classification differentiates broad potential habitats based on abiotic surrogates. *In situ* data of substrate surrounding Santa Catalina Island provided by the National Oceanic and atmospheric Administration (NOAA) were then overlaid onto the unsupervised classification to identify spatial associations between substrate type and abiotic classes using the “multi-values to points” tool in ArcGIS Pro v. 2.3.0 (19). Those substrate data were collected in 1975 using grab samples, which were then classified based on material and grain size. To simplify the analysis, the thematic scale of those data was coarsened by regrouping similar classes together. For example, in NOAA’s classification, sand classes were differentiated based on color. We grouped all sand colors into one class of sand for our analysis. These substrate data provided environmental context to the unsupervised classification. Classes were then compared with the occurrences of giant kelp and devil weed (see below).

The second approach to habitat mapping enables the prediction of the potential habitat distribution of *M. pyrifera* and *S. horneri* around Santa Catalina Island based on their respective association with the environment. *Sargassum horneri* occurrences were obtained from Marks *et al.* (2015) (8), who made their dataset available through a long-term repository (20). This dataset includes compiled records of *S. horneri* observations recorded during ecological surveys by trained observers, from 2003 to 2015. The dataset was filtered to keep only the occurrences surrounding Santa Catalina Island and to remove duplicate locations, resulting in 349 occurrences. However, due to the presence of a littoral gap in the abiotic data (see discussion), only 57 occurrences were overlapping the abiotic data. Giant kelp occurrences were provided by the California Department of Fish and Wildlife, which has performed aerial surveys most years since 2002. The main survey data are multispectral imagery (red, green, blue, near-infrared) at 0.3 m resolution and target both submerged and exposed kelp beds. Giant kelp coverage is extracted from the imagery and provided yearly as polygons representing the extent and distribution of kelp. In order to ensure comparability with *S. horneri* data, we compiled only the *M. pyrifera* coverage for the years 2003 to 2015. A set of random points was then created using ArcGIS Pro in order to maximize within-polygon coverage while keeping a minimum distance of 50 m between sample points to prevent autocorrelation. A total of 347 point samples were created as a result. Of those points, 38 overlapped the abiotic data. The point occurrence data of each species were then used in MaxEnt to produce two maximum entropy habitat distribution models – one for each species. Maximum entropy modeling is a machine learning method for species distribution modeling that is often used to model geographic distributions of suitable habitats from presence-only data (21). Models were cross-validated 30 times each with the final model representing the average of those 30 models. Jackknife measures of variable importance and response curves were also generated to identify which environmental variables drive the species distribution and contribute to the models. The area under the curve (AUC) of the receiver operating characteristic curve (ROC) was used as a performance measure for the MaxEnt models. The ROC illustrates the discrimination ability of a model by plotting the true positive rate against the false positive rate at various thresholds. The two habitat distribution models were overlaid to identify the areas corresponding to suitable habitat for both

species. The habitat suitabilities for each pixel of the resultant maps were multiplied by each other to develop a map indicating the probability of the species co-occurring at every pixel, on a scale from zero to one. Additionally, a binary dataset was developed for both species, classifying any pixel that had a suitability greater than 50% as highly suitable habitat, as commonly done in habitat suitability modeling (22). The two binary datasets were then added together in ArcGIS to quantify the area that corresponded with an overlap of suitable habitat.

Results

The analysis of correlation among variables identified two pairs of covarying variables: slope and rugosity, and TPI (20 m) and TPI (50 m) (Table 1). Thus, the selection of variables for the remaining analyses was bathymetry, easternness, northernness, slope, TPI (20 m), and TPI (250 m).

	Bathymetry	Easternness	Northernness	Rugosity	Slope	TPI20	TPI50
Easternness	-0.002						
Northernness	0.000	0.001					
Rugosity	-0.100	0.003	0.002				
Slope	-0.110	0.003	0.003	0.984			
TPI20	0.123	0.001	0.001	0.329	0.299		
TPI50	0.238	0.002	0.002	0.444	0.429	0.712	
TPI250	0.395	0.003	0.002	0.541	0.556	0.358	0.628

Table 1. Correlation matrix between the different variables available for the analysis. Slope and rugosity, and TPI20 and TPI50, were highly correlated ($r \geq 0.700$). Rugosity and TPI50 were thus not used in the analyses.

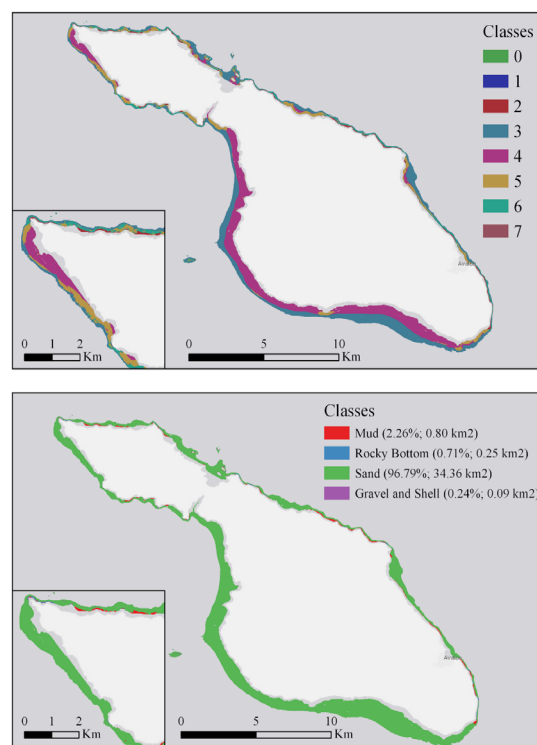


Figure 3. Results of the unsupervised classification, which defined eight potential habitat classes. Those classes were spatially compared to ground-truth data to interpret them as four substrate classes. The insets show the northernmost part of the island.

Using a k-means classification, we created a map of undefined potential habitat classes that are statistically distinct. Eight distinct classes were determined for the benthic area surrounding Santa Catalina Island (Fig. 3). This habitat map based on abiotic surrogates was then validated with *in*

situ substrate data from NOAA to be regrouped into classes representing the actual habitats. This resulted in the eight undefined but distinct classes being regrouped into four defined classes with unique substrate types: sand, mud, rocky bottom, and gravel and shells (Fig. 3). The overall accuracy of this map was relatively low yet acceptable at 58%. Most of the area was identified as being sand (96.79%), although the shallowest areas along the coastline were often predicted to be either mud (2.26% of the surveyed area), rocky bottom (0.71% of the area) or gravel and shells (0.24%), potentially providing insights into the benthic composition in the littoral gap (see discussion). Most of the *M. pyrifera* (78.9%) and *S. horneri* (52.6%) samples were found in sandy areas. The remaining giant kelp samples were found on mud and rocky bottom (10.5% each). The remaining devil weed samples were found on mud (26.3%), rocky bottom (15.8%), and gravel and shells (5.3%).

For the application of the supervised approach to habitat mapping to produce habitat suitability maps of giant kelp and devil weed, both models performed well: the mean AUC (of the 30 models used for cross-validation) for the giant kelp model was 0.89 and the mean AUC for the devil weed model was 0.90, which are considered excellent (models with an AUC greater than 0.50 are better than random and those with AUC greater than 0.75 are considered to provide a useful amount of discrimination; see 23).

	Percent Contribution		Environmental Preference	
	<i>M. pyrifera</i>	<i>S. horneri</i>	<i>M. pyrifera</i>	<i>S. horneri</i>
Bathymetry	78.2%	2.6%	Shallower waters	Shallower waters
Easterness	1.8%	2.5%	No discernable pattern	No discernable pattern
Northernness	4.3%	1.5%	No discernable pattern	No discernable pattern
Slope	6.6%	4.3%	Higher slopes	Higher slopes
TPI20	0.0%	0.0%	Elevated seafloor features	No discernable pattern
TPI250	9.1%	88.9%	Elevated seafloor features	Elevated seafloor features

Table 2. Results from the MaxEnt models, including the average percent contribution of each variable to the 30 models (*i.e.*, cross-validation) for both species, and the environmental characteristics more suitable for each species' settlement.

In terms of variable importance, which corresponds to the percentage contribution of each variable to the models (and thus in explaining species distribution), giant kelp was largely driven by depth, a likely proxy of light availability, and to a lesser extent slope (Table 2). Devil weed habitats were largely driven by broad-scale topographic features (*i.e.*, TPI-250 m) and slope (Table 2). The analysis of species-environment relationships highlights many similarities between *S. horneri* and *M. pyrifera*: both species were more readily found in shallow waters, on higher slopes, and in elevated areas of higher TPI values.

Fig. 4 shows the habitat suitability maps that were produced with MaxEnt for giant kelp and Fig. 5 shows the habitat suitability models for devil weed. Those maps show that 10.27% (about 3.65 km²) of the sampled area surrounding Santa Catalina Island (in waters shallower than 55 m) is suitable habitat for *M. pyrifera* and 7.37% (about 2.62 km²) is suitable habitat for *S. horneri*. Of the area suitable for giant kelp, 33.56% was also suitable for devil weed (Fig. 6). Most of the area of high probability of overlap was located in the northern section of the island and the northeastern side (Fig. 6). The southwestern side of the island was suitable for giant kelp but not for devil weed (Fig. 4 to 6).

Discussion

Both types of habitat maps produced in our study show that the two species of macroalgae seem to thrive in similar environmental conditions. Our map of potential habitats (Fig. 3) suggest that they mainly settle in sandy environments and that *S. horneri* might be more of a habitat generalist than *M. pyrifera* by inhabiting other types of benthic environment like mud and rocky bottoms. Those results are to be interpreted with care as the overall accuracy of this map was not very high, and the area covered by the abiotic data may not be fully representative of

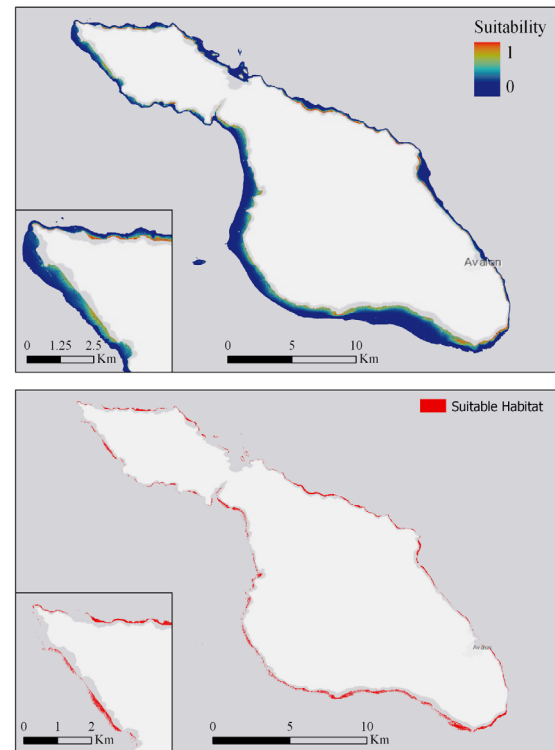


Figure 4. Maximum entropy model for *M. pyrifera* identifying the suitable habitat for this species. The bottom panel represents the areas with a suitability higher than 0.50 (considered highly suitable (22)). The insets show the northernmost part of the island.

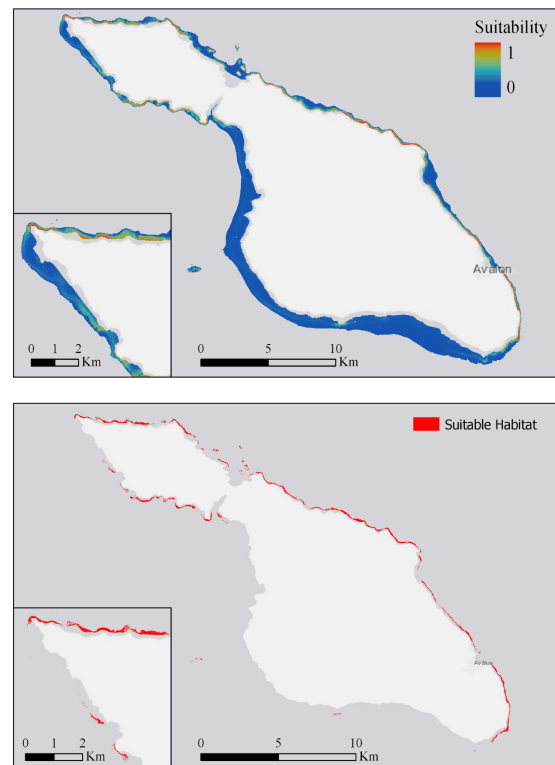


Figure 5. Maximum entropy model for *S. horneri* identifying the suitable habitat for this species. The bottom panel represents the areas with a suitability higher than 0.50 (considered highly suitable (22)). The insets show the northernmost part of the island.

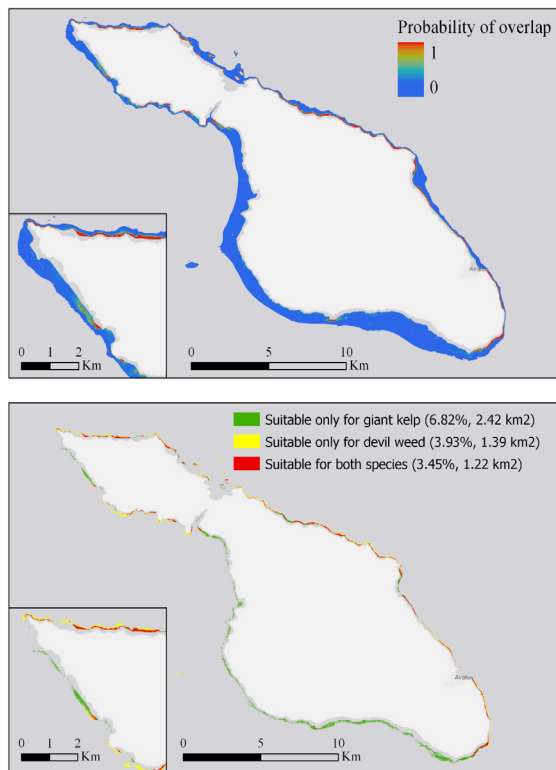


Figure 6. Probability that both species overlap, based on the multiplication of their habitat suitability. The insets show the northernmost part of the island.

the range of environmental characteristics preferred by the two species (see discussion on littoral gap below). Our species distribution models show that about a third of the suitable habitat for native giant kelp off the coast of Santa Catalina Island is also suitable for non-native devil weed, thus highlighting a high potential of invasion. Based on the generated map, giant kelp seemed to thrive in northward and southward facing areas, which is consistent with seasonal current patterns in the area (24). By positioning itself on a slope parallel with the currents, *M. pyrifera* places itself in the optimal position for gathering upwelling nutrients as the seafloor grows shallower. However, we note that El Niño events shift patterns of upwelling and introduce warmer waters into the ecosystem (25). This, in turn, can reduce native kelp growth and allow *S. horneri* to thrive, and highlights the need for adaptive management techniques (12, 26). More research is needed to understand the more consistent distribution across slopes for *S. horneri*.

Our maps can be paired with current research that examines strategies for patterns of *S. horneri* removal as it can provide additional knowledge about its invasive capabilities. There has been little work done on identifying the ecological variables that allow *S. horneri* to invade kelp forests. Consequently, management practices currently rely on managing giant kelp populations and targeted removal of devil weed. Understanding the spatial dynamics of *S. horneri* and *M. pyrifera* habitat is a necessary first step to identifying these features and provides decision-makers with the necessary tools to reduce further loss of kelp forests via *S. horneri* occupation. The efficiency of *S. horneri* removal is based on several factors, including time of year and methods (12). Understanding the temporal dynamics of these species and producing temporally-explicit species distribution models might offer new insights on management and removal approaches. Removal of new generations is more effective in autumn when *S. horneri* are juveniles and intra-specific competition is at its highest than in winter when they have matured to adulthood (12). The most appropriate technique for removal of the non-native species is dependent upon the budget and staff available. Focusing on areas with a lower chance of *S. horneri* re-establishing and selecting efficient time periods and methods of removal (e.g., suction) will maximize results. Marks *et al.* (2017) demonstrated that light availability is a contributing factor to *S. horneri*'s invasion fitness (12). It is therefore possible that

S. horneri's association with more positive topographical features (i.e., positive TPI) might be a mechanism that allows it to move higher in the water column to absorb more light. *Macrocystis pyrifera*'s size does not require such elevated features since its growth into the water column places it in a more optimal position for photosynthesis. In addition, giant kelp is more capable of succeeding with low light situations (26). The alga is largely self-shading, with its leaves reducing the light available for photosynthesis at 1 m depth (26). The sub-canopy of the plant can thrive from high-intensity sunflecks that occur when wave-action shifts the plant in the water column. Thus, shading could potentially be used as a more cost-effective, but still targeted, approach to *S. horneri* removal.

The results of this study provide insight into the interactions between kelp forest ecosystems and devil weed in the waters surrounding California's other Channel Islands. The relationships that were established between benthic terrain and the distribution of both species may be used to predict suitable habitats for *S. horneri* and their overlap with suitable habitats for *M. pyrifera* around other Channel Islands and in coastal waters. Such habitat suitability maps could inform monitoring on other islands and enable the eradication of devil weed as soon as it appears in other locations. Our study demonstrated the potential for habitat mapping and species distribution modeling to be used as spatial decision-support tools for Santa Catalina Island and in other areas that require similar map products. However, we note that the extent of our study area was limited by the presence of a littoral gap in the dataset used to produce the maps and models. An example of that littoral gap can be seen in Fig. 2, where data coverage does not reach the coastline – land is represented in white in the figure. In remote sensing, a littoral gap, or white ribbon, refers to the shallow areas where it is difficult to collect data (26): data collected from acoustic remote sensing cannot be collected due to the inability of the surveying vessels to navigate in shallow waters, and terrestrial remote sensing techniques are often not adequate to study the seafloor (27). In our study area, the littoral gap was varying from 35 to 700 m from the coast. The latter measurement reflects Isthmus Cove, a wide and shallow bay on the northeast side of the island. Given how close to shore giant kelp populations are commonly found (Fig. 1C), this littoral gap prevented our models from capturing the entire suitable habitat for giant kelp. A more comprehensive coverage of kelp habitat would improve the accuracy of the predicted distributions around Santa Catalina Island and elsewhere. Future studies could look into using bathymetric LiDAR, or laser altimetry, to fill the littoral gap and collect seafloor data close to shore (28), thus covering the entire range of kelp habitats. The temporal gap between the substrate data collected in 1975, the abiotic data collected in 2008, and the biotic data compiled from 2003 to 2015 is another limitation to consider, as habitat and environmental conditions might have changed to some extent over this timeframe. While remote sensing techniques can fill the littoral gap, *in situ* data collected by grab samples or visual observation by divers or towed video data would also result in more informed maps. Given Santa Catalina Island's active diving community, it would be feasible to employ citizen science as a method to obtain more robust validation data, for instance to collect different types of seafloor data and capture images to document and later interpret benthic composition.

Conclusion

Non-native algae such as *S. horneri* must compete with native algae species for space, understory, and resources such as light (11). Studies have shown reduced recruitment and survivorship of non-native sargassum species in communities containing native algae that cover ground space such as turfing algae and tall canopy-forming kelp (29). Thus, if removal efforts were targeted toward areas where *M. pyrifera* is still prevalent, the native algae species could re-exert biotic control on the ecosystem and reduce the threat of *S. horneri* re-establishment (12). This study identified areas of suitable habitat for giant kelp that are also likely to eventually be invaded by devil weed and is a vital first step to identifying areas that would be most effective for removal.

Acknowledgements

Thanks are due to Dr. Lindsay Marks at the University of California, Santa Barbara for providing biological data on *Sargassum horneri* presence.

References

1. Gaitan-Espitia JD, Hancock JR, Padilla-Gamino JL, Riverest EB, & Blanchette CA. Interactive effects of elevated temperature and pCO₂ on early-life-history stages of the giant kelp *Macrocystis pyrifera*. *J Exp Mar Bio Ecol*. 2014; 457, 51-58.
2. Behringer D. Primary Producers. Lecture presented in Marine Ecological Processes at University of Florida, Gainesville, FL. 2018 August 30.
3. Graham MH. Effects of Local Deforestation on the Diversity and Structure of Southern California Giant Kelp Forest Food Webs. *Ecosystems*. 2004; 7(4), 341-357.
4. Johnson A. California's fertile forests. *The Island Naturalist*. 2013; 28.
5. Catalina Chamber. Economic and fiscal impacts and profile of 2016 Catalina Island visitor final report. 2017 June.
6. Catton C, Rogers-Bennett L, Amrhein A. Perfect storm decimates northern California kelp forests. *CDFW Marine Management News*. 2016 March 30.
7. Schuman C. California Department of Fish and Wildlife aerial kelp survey results, 1989-2016. 2016. Published raw data.
8. Marks LM, Salinas-Ruiz P, Reed DC, Holbrook SJ, Culver CA, *et al*. Range expansion of a non-native, invasive macroalga *Sargassum horneri* (Turner) C. Agardh, 1820 in the eastern Pacific. *Bioinvasions Records* 2015; 4(4): 243-248.
9. Ginther S, Steele M. Limited recruitment of an ecologically and economically important fish, *Paralabrax clathratus*, to an invasive alga. *Mar Ecol Prog Ser*. 2018 August 23; 602: 213-224.
10. Miller KA, Aguilar-Rosas LE, & Pedroche FF. A review of non-native seaweeds from California, Mexico. *Hidrobiológica*. 2011; 365-379.
11. Hubbard, S. Exploring California's Marine Protected Areas: Blue Cavern Onshore State Marine Conservation Area and Blue Cavern Offshore State Marine Conservation Area. 2018 May 22.
12. Marks L, Reed D, Obaza A. Assessment of control methods for the invasive seaweed *Sargassum horneri* in California, USA. *Manag Biol Invasion*. 2017; 8(2), 205-213.
13. Brown CJ, Sameoto JA, Smith SJ. Multiple methods, maps, and management applications: Purpose made seafloor maps in support of ocean management. *J. Sea Res*. 2012; 72.
14. Seafloor Mapping Lab. SFML data library. California State University, Monterey Bay. 2008.
15. Perissinotto R, McQuaid CD. Deep occurrence of the giant kelp *Macrocystis laevis* in the Southern Ocean. *Mar. Ecol. Prog. Ser.* 1992; 81: 89-95.
16. Lecours V, Devillers R, Simms AE, Lucieer VL, Brown CJ. Towards a framework for terrain attribute selection in environmental studies. *Environ Model Softw*. 2017; 89, 19-30.
17. Lindsay JB. Whitebox GAT: A case study in geomorphometric analysis. *Comput Geosci*. 2016; 95: 75-84.
18. Patel VR, Mehta RG. Impact of Outlier Removal and Normalization Approach in Modified k-Means Clustering Algorithm. *IJCSI*. 2011; 8(5).
19. H09497: NOS Hydrographic Survey, Santa Catalina Island, California, 1975-03-30, available at <https://catalog.data.gov/dataset/h09497-nos-hydrographic-survey-santa-catalina-island-california-1975-03-30>. Last accessed on 02/20/2019.
20. Marks LM, Salinas-Ruiz, P, Reed DC. Records of *Sargassum horneri* occurrence in the eastern Pacific. *Environmental Data Initiative* 2015, <https://doi.org/10.6073/pasta/63012c4e436214239ebed11ee57cbe03>. Dataset last accessed on 02/20/2019.
21. Phillips S, Anderson R, Schapire R. Maximum entropy modeling of species geographic distributions. *Ecol Model*. 2006 January; 190(3-4): 231-259.
22. Zohmann M, Pennerstorfer J, Nopp-Mayr U. Modelling habitat suitability for alpine rock ptarmigan (*Lagopus muta helvetica*) combining object-based classification of IKONOS imagery and Habitat Suitability Index modelling. *Ecol. Model*. 2013; 254: 22-32.
23. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, *et al*. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 2006; 29: 129-151.
24. Britannica. (2011, November 06). California Current.
25. Zimmerman RC, Robertson DL. Effects of El Niño on local hydrography and growth of the giant kelp, *Macrocystis pyrifera*, at Santa Catalina Island, California. *Limnol. Oceanogr*. 1985; 30(6), 1298-1302.
26. Gerard VA. 1984. The light environment of a giant kelp forest: influence of *Macrocystis pyrifera* on spatial and temporal variability. *Mar. Biol*. 1984; 84: 189-95.
27. Lecours V, Dolan MFJ, Micallef A, & Lucieer VL. A review of marine geomorphometry, the quantitative study of the seafloor. *Hydrol. Earth Syst. Sci*. 2016; 20: 3207-3244.
28. Copeland A, Edinger E, Devillers R, Bell T, Leblanc, P, & Wroblewski, J. Marine habitat mapping in support of Marine Protected Area management in a subarctic fjord: Gilbert Bay, Labrador, Canada. *J Coast Conserv*. 2011; 17(2), 225-237.
29. Britton-Simmons KH. (2006). Functional group diversity, resource preemption and the genesis of invasion resistance in a community of marine algae. *Oikos*. 2006; 113(3): 395-401.

Research Article

¹McGill University, Montreal,
QC, Canada

Keywords

High Arctic runway, Soil compaction, Periglacial environment, Hydrogeomorphology

Email Correspondence

tabatha.rahman@mail.mcgill.ca

Tabatha Rahman¹, Jeremiah Lee¹

Physical Responses to Anthropogenic Disturbance on a Runway in the Canadian High Arctic

Abstract

Background: The McGill Arctic Research Station (MARS) runway on Axel Heiberg Island, Nunavut, has been subject to intense compaction from aircrafts since the 1960's, causing alterations to the landscape across time and space. This study investigated the thermal, hydrological and physical responses of repeated aircraft compaction of the runway to characterize its vulnerability to land use, and the effects of long-term and repeated compaction.

Methods: In late July and early August 2018, the island's summer thaw season, topographic, soil, hydrological and frost table data were collected along four transects across the MARS runway.

Results: Topography and effective porosity differences caused by compaction were found to affect soil moisture contents, leading to the observed differential frost heaving and insulation properties of soil across the runway. Soil was found to be mostly silt on and off disturbed areas, indicating that compaction does not affect grain-size but rather pore space and soil bulk density. The frost table mapping suggested statistically significant variations in depth of the frost table across undisturbed, disturbed, and indirectly disturbed areas, showing that compression from aircrafts has both direct and indirect spatial impacts on the hydrogeomorphic system. Furthermore, this research examined possible solutions to mitigate thaw consolidation of the runway.

Limitations: The method of probing used to determine frost table depths introduced significant error to the data. Probing discrepancies arose between people probing and between sampling days, as techniques differed and/or improved. Future studies should consider using electrical resistance tomography to map the frost table, as this would eliminate inconsistencies. Furthermore, while pore size distributions were inferred based on grain-size and extent of compaction, subsequent studies should consider a quantitative approach to pore space analysis.

Conclusion: This study suggests that aircraft travel to the remote McGill Arctic Research Station causes spatially and temporally significant changes in the local hydrogeomorphology, especially in fine-grained and wet, frost-susceptible soils. Thaw consolidation, which results indicate is caused by the direct and indirect effects of soil compaction, compromises the prolonged use of the runway.

Introduction

The runway at the McGill Arctic Research Station (MARS), near Expedition Fjord on Axel Heiberg Island, Nunavut, has undergone significant physical disturbance since Expedition Fjord gained scientific interest following Fritz Müller's first expedition in 1959 (1). Small aircrafts have repeatedly compacted the MARS runway, and the ice-wedge polygons found on the runway have been infilled over the years to counter the effects of subsidence, as shown on Fig. 1. As thaw consolidation of the ice-wedges continues, an understanding of the hydrogeomorphic system in this dynamic periglacial environment becomes crucial for the prolonged use of the runway.



Fig 1. Transect 3 with its polygon troughs infilled with gravel and rounded stones.

The MARS runway has summer air temperatures averaging 5°C (2) and is considered a polar desert (3). During the late spring and summer, its mineral soils are moist and even saturated, as soil water and groundwater are restricted to the thin portion of thawed ground above the perennially frozen soil. This is referred to as permafrost. Soils in polar regions are subject to intense mechanical weathering causing patterned ground, such as polygonal terrain or ice-wedge polygons. Considering water expands by 9% when freezing, and freeze-thaw cycles cause rock and soil weathering, erosion, and displacement (4). Thus, frost action and frost heaving are leading processes in landscape alteration of this environment.

Soil and water interactions (i.e. hydrogeomorphic processes) are of major importance to this study (5). Hydrological properties, such as hydraulic conductivity, moisture content, drainage, and water retention capacity of the soil affect the magnitude of frost heaving and the formation of ice-wedge polygons. Ice formation within the active layer, also known as the cryostructure, affects the ground's susceptibility to subsidence, which is a safety concern for the MARS' runway.

Through analyses of the frost table depth, topography, soil moisture content and particle size distribution of soil, this research sought to spatially and temporally define the thermal, hydrological and physical responses of repeated aircraft compaction on and around the runway of the McGill Arctic Research Station. This paper provides a qualitative and quantitative overview of the responses of repeated compaction of the runway, which could be used for future mitigation efforts to ensure the continued use of this Arctic tundra runway.

Methods

Field Methods

The runway (79°24'55.1" N; 090°45'51.0" W) lies in an unglaciated valley with significant soil development. Its gentle slope allows for water drainage towards the areas of lowest elevation: Colour Lake and the marsh (Fig. 2). Low-lying polygonal troughs (i.e. ice-wedges) act as surface preferential flow paths, further supporting drainage towards the lake and marsh. The ice-wedge polygons are present on and around the runway, displayed in Fig. 1 and Fig. 3.

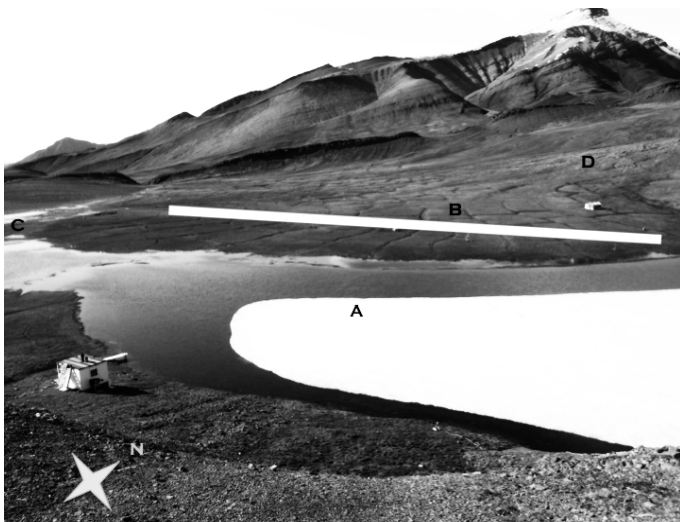


Fig. 2. The topographic profile of the runway, from SW to NE. Transect 1 is the left-most, followed by transects 2, 3 and 4.

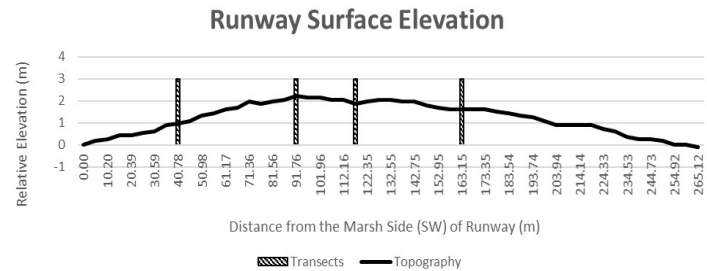


Fig. 3. View of the runway, here shown by the white arrow line at the center, from MARS. A is Colour Lake (partially covered by ice), B is the polygonal terrain, C (at the very left) is the marsh and D is the alluvial fan.

Field data was collected in late July and early August 2018, during the local summer thaw period. The length of the runway is 265 m Southwest to Northeast and 10.5 m wide. Four 35 m transects, henceforth named T1, T2, T3 and T4, were established perpendicular to the length of the runway to incorporate the variability in the conditions along its length. Among the four transects, only T3 and T4 were defined by polygonal terrain, while T1 and T2 had desiccation cracks. T3 had some gravel and a large cobble-filled polygon trough, while T4 had polygon troughs filled with Styrofoam and gravel. The difference in surface patterns on each transect illustrates the small-scale hydrogeomorphic differences across the runway.

A 1 m resolution topographic profile of each transect was built using a level string at an arbitrary height above each transect, from which the soil-string distance was measured at every meter along the transect.

A 5 m resolution topographic profile of the length of the runway was also built (Fig. 2). Two poles, held vertically and interconnected by a 5m rope, were transported across the runway, while the horizontal distance between them was measured. Using a Brunten (Forestry Suppliers, Inc), the angle formed by the rope from the horizontal was measured. Basic trigonometry then allowed for the changes in elevation to be calculated.

Each transect was subdivided into three sections: undisturbed, disturbed, and indirectly disturbed. The undisturbed sections comprised points 1 to 8 along the transects, or the first eight meters upslope (North) of the runway. The disturbed regions were the runway portions of the transects, from point 9 to point 21. The indirectly disturbed sections were the section downslope (South) of the runway, from points 22 to 36. It was hypothesized that this third region was indirectly affected by aircrafts through disturbances in the hydrological and thermal regimes radiating from the disturbed locations. This was tested by comparing the mean frost table depth in each section.

The frost table depth was measured using a 1.5 m stainless steel rod. Although depths were measured to the nearest centimeter, a significant error was introduced due to the high presence of rocks and compacted soil material. When it was assumed that the rod hit a rock rather than frozen ground or when there was a discrepancy of more than 3 cm from the previous day's probing depth, four additional data points within a 0.4 m radius were taken. The deepest data point was kept and considered to be the frost table depth. If all data points were 3 cm closer to the surface than past probe depths, the data was excluded from analysis, as bottom-up freeze up is unlikely at that time of year. Probing discrepancies also arose between people probing and between days, as techniques improved. Future studies should consider using electrical resistance tomography to map the frost table, as this method would eliminate inconsistencies and would be more precise, albeit more expensive.

For particle-size and moisture content analyses, two soil samples were collected per transect on either side of the runway (upslope and downslope). The purpose of these samples was to compare moisture contents upslope and downslope from the disturbance to determine if compaction affected the hydrological system.

Laboratory Methods

The gravimetric field moisture content was calculated in lab using differential weighing of dry and moist soil, to find the percentage air-dry mass of soil. All soil samples' grain size distributions were tested using Kroetsch et al. (2007) (6) hydrometer method to separate silt, sand and clay. Soil solutions were then poured into 2mm and 63µm sieves to separate larger particles such as gravels and sand.

Results

Transect Topography and Frost Table Depth

Using an ANOVA F-test with a completely randomized design to compare 3 treatment means and $\alpha = 0.05$, T1, T2 and T4 were found to have statistically different frost table depth between their undisturbed, disturbed and indirectly disturbed sections, and only the third transect's frost table was found to be statistically unchanging across. However, this finding is believed to be caused by an outlier. T1, T2 and T4's three sections had at least two of their sections' mean frost table depth that differed from one another, suggesting that the anthropogenic disturbance to the runway caused significant changes to the frost table depth. As shown in Table 1, T1 and T4 saw a general deepening of the frost table going downslope. T1's frost table decreased by 12.5 cm from the undisturbed to the disturbed area, and stayed relatively constant from the disturbed to the indirectly disturbed section. T2 saw a local depression in its frost table on the runway, as there was a decrease of 19.2 cm from the undisturbed section to the disturbed, and a 20.0 cm increase in the frost table from the disturbed to the indirectly disturbed. T3's variations in frost table depths were less than 5.6 cm and statistically insignificant. T4's frost table descended by 6.9 cm, and then again by 2.1 cm.

	Mean depth of the frost table			
	T1	T2	T3	T4
Undisturbed	40.8	54.8	53.0	60.3
Disturbed	53.3	35.6	50.1	67.2
Indirectly Disturbed	53.1	55.6	55.8	69.3

Table 1. Mean frost table depth (cm) within three subsections of each transect of the runway.

Grain Size Distribution

Grain size distribution was used to infer hydrological and thermal properties of the soil. Silt was determined to be the main component of the soils on and around the runway. Soils analysed from a nearby undisturbed polygon trough and top were compared to the disturbed runway soil. The grain size distributions were consistent to the ones on the runway, suggesting compaction does not affect grain size.

Soil Moisture Content

As shown on Fig. 4, T1 and T3 saw a decrease of moisture content downslope. The difference in moisture content between the upslope and downslope edges of the runway suggests that compaction on T3 alters the ability of water to flow in the subsurface. Evidence of pooling of water at the upslope edge of the runway on T3 further corroborates this theory, as compaction of soil acts as a physical barrier to groundwater flow. Compacted soils have less pore space in which water may travel in and cause water accumulation, in turn causing subsidence.

Gravimetric Field Moisture Content

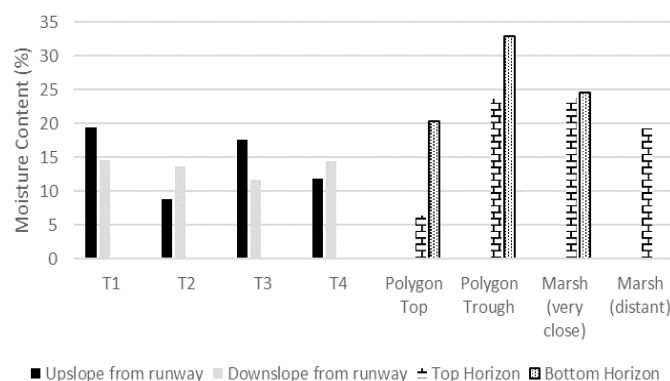


Fig. 4. Gravimetric field moisture content of soils and around the runway at MARS. Samples were taken upslope (e.g. North) of the runway and across the runway (e.g. South) along each transect.

The most prominent difference in moisture contents was found between the topmost horizons of the polygon peak and trough. This is most likely due to the topographic difference between the two, as the trough lies closer to the water table than the peak.

Discussion

Polar soils are most susceptible to disturbance in late spring, while thawing of the active layer has begun and there is a thin layer of thawed and saturated soil above the frozen ground (7). Much of the aircraft traffic on the MARS runway happens during this crucial time, as access to the research station is restricted during winter.

Aircrafts' usage of the runway at Expedition Fjord certainly disrupts the natural equilibrium of the land, but the extent to which it does so depends on the frequency, intensity, scale and timing of these disturbances (7). The intensity of impact between the aircraft and the runway is relatively constant. Only helicopters and bush planes, including Twin Otters with empty weights varying between 2653 kg and 3363 kg (8), can land at the MARS runway. The frequency is restricted to less than 20 landings and takeoffs per year, over several decades. This repeated disturbance does not allow the system to reach a new steady-state, but rather creates a dynamic, evolving system.

One of the most important responses to the runway's land use on a safety point of view is the subsidence of ice-wedge polygons on the runway. An uneven surface is a safety hazard for aircrafts, and remedying this problem is necessary for the prolonged use of the runway. Gravel, stones and Styrofoam have been inserted into subsiding polygon troughs to prevent further thawing and to flatten the surface of the runway (Fig. 1). Kevan et al. (1995) (6) studied vehicle tracks' effects on soil and vegetation at Hazen Camp, on Ellesmere Island (Nunavut), and observed that ice-wedge polygons deepened and widened where tracks were present. Their paper argued that subsidence in wet areas was the effect of direct compression from vehicle tracks. Ice-wedge polygon subsidence is indeed related to the thermal and hydrological systems. Usually, the amount of water in frozen soil is superior to that which can be held in thawed soil (4). When thawing begins, sediments settle as the excess water drains, causing thaw consolidation. At the MARS runway, as subsurface water flows encounter the compacted soil of the runway as they flow downslope, there is a damming effect on the North (upslope) side of the runway in polygonal terrain. The preferential flow paths within the polygonal troughs are assumed to have been disturbed and suppressed by the compacted soil of the runway, causing water to pool and thus causing thaw consolidation, as shown on Fig. 5. This is an example of the radiating effects of the disturbance, as this feedback effect is not located on the runway itself, but rather on ground that is not directly used by the planes. This reinforces the idea that in the Arctic, disturbances have rapid cascading feedback effects (9) that tend to extend through space and time.



Fig. 5. The white box shows the local depression that formed on the NW side of the runway near T4, as a response to the disturbed subsurface water flow.

Variations in the active layer and frost table depths are other responses to the disturbance by aircrafts. As shown in Table 1, there is a statistically significant difference in frost table depth across transects 1, 2, and 4. The active layer depth depends on thermal systems within soils, as thawing is enabled through the thermal conductivity of soils and soil-waters. Forbes (10) showed that physical disturbances cause greater variations in diurnal and seasonal soil temperatures than undisturbed soils. Greater variations imply increased potential for freeze-thaw cycles, imposing greater mechanical strain on soils. In depths greater than 5.0 cm – 10.0 cm, only annual freeze thaw cycles take place (4), suggesting that most of the frost action (e.g. frost heave) takes place in the upper 10 cm of soil. As a result, most of the preferential sorting occurs within this topmost layer of the ground, which coincides with the layer that is subject to the greatest compaction.

Some soils are prone to stronger compaction feedback effects. Studies suggest that the ability of soils to compact is related to numerous factors, notably the cohesive force between particles (11, 12). Compaction of Arctic soils by vehicles has shown to result in shifts in ground albedo, soil moisture, pH, active layer development, soil temperatures, vascular plant biomass, total species richness and mineral nutrition (10). High Arctic thermal and plant nutrient cycling regimes have shown to be affected for decades after a single vehicle pass (10), demonstrating the extended temporal effect of vehicular use on Arctic soils. Lowery and Schuler (1991) studied silt loam, similar in particle-size to the MARS runway soils, and concluded that compaction by heavy farm equipment over one year lowers plant height for over 4 years (13). Vegetation differences were clear between the runway and the areas adjacent to it; the runway had extremely few plants growing on it. Because plants are important sources of insulation, since they create an organic layer that buffers air temperatures from the soil, the removal of plants further affects the thermal regime of the runway.

Furthermore, compaction affects water retention curves and hydraulic conductivity, as porosity is altered in compacted soils (14). In the event of stagnation of water during freeze-up, thaw consolidation is more likely to happen upon thawing the following spring, as large amounts of water within the soil matrix are released. The soil on the MARS runway would therefore benefit from both high porosity, to allow for drainage, and from fine pores, to allow moisture movement under dry conditions. As large clasts are less likely to see a reduction in effective porosity with compaction, infilling subsiding troughs on the runway with gravel and small stones would allow for pore spaces to be maintained, even under immense compaction.

Compaction also leads to a decrease in the insulative properties of soil, as pore spaces are reduced and gas, which acts a buffer between surface air temperatures and soil temperatures, is pushed outside of the soil matrix. While the results of this study do not suggest compaction impacts soil texture, future studies should look at pore size distributions on and off the runway, as compaction does affect the soil's ability to store and transmit

water and heat (15, 16) by changing the effective porosity of soils, which in turn affects plant growth and periglacial processes.

Compaction, combined with the increased thermal conductivity of soils due to the removal of vegetation, increases the runway's susceptibility to thermal erosion. Shilts (1978) found a thicker active layer in bare ground and fine-grained soils, compared to vegetated, coarse grained soil (17). This finding is consistent with T1 and T4's frost table depth on the runway, and suggests that the fine-grained bare soils of the runway are not good insulators. However, a deeper thawing was not present on T2, as the frost table was closer to the surface. T2's location in relation to the other transects indicates that it is at a high point (Fig. 2), allowing for relatively better drainage. Because thermal conductivity of soil rises with increasing moisture content (18), it was hypothesized that this transect was dryer than the others, as its frost table was shallower than the other transects. However, the results from the gravimetric water content tests do not fully corroborate this.

The scale of the disturbance response is also influenced by the topography of the landscape, as it affects thermal and hydrologic cycles within the soil. Microtopography determines subsurface and surface water flow velocities and determines some of the sources of water made available to the system. In active layer hydrology, flowing subsurface water has been known to have a warming effect on the underlying frozen ground (18). For example, water tracks in the Antarctic polar desert had active layers twice as thick as adjacent areas (19). However, for a warming effect to take place, the residence time of water must be long enough to allow for advection (14). The shallow slope of the MARS runway enabled advection from the water to the surrounding soil because surface and subsurface water flow had negligible velocity. Advection may be involved in the thaw consolidation of polygon troughs and of the upslope boundary of the runway near T3 (Fig. 5). However, the source of subsurface water also influences its ability to warm the active layer. While snow or ground ice melt releases water with little available heat, rain events deliver more heat to the ground. As rain is expected to be the dominant form of precipitation in the Arctic by the 2080 (20), there will be considerably more heat entering the soil system of the runway at MARS, causing further subsidence.

Conclusion

Scientific research at Expedition Fjord has caused landscape and hydrological alterations of the MARS runway. As the scale of impact is not equivalent to the scale of response, mitigation measures should be taken to prevent further thaw consolidation and subsidence, yet solutions are limited. A static equilibrium recovery cannot be envisioned for the runway, as disturbances are repeated from year to year. Becker and Pollard (2016) found that a High Arctic airstrip, unused for the past 60 years, still has not reached its pre-disturbance conditions, but rather has moved towards an ecological succession with a new stable-state community (21). Considering the ongoing repeated disturbance events on the MARS runway, this study suggests that disturbance responses can be mitigated by allowing the thawed active layer to transport water and by increasing the insulative properties of the soil.

Airports and landing strips across northern latitudes have faced similar problems with frost heave and thaw settlement; the main problem is the obstruction of groundwater flow paths. Major Northern airports that welcome large freight and passenger aircrafts and that have, as a result, more regulated safety measures, have paved asphalt runways, with excavated material beneath them filled with non-frost susceptible structural fill. However, this is not a viable option for smaller airstrips, where the use of gravel and rocks as infill (7), similar to MARS' runway, is more common (4). Future studies of the MARS runway should investigate subsurface water flow paths, which would provide insight on how to potentially redirect water flow to artificial preferential flow paths, thus diminishing subsidence on crucial locations on the runway.

Acknowledgements

This research was funded by McGill University's Dr. Olav Loken Undergraduate Geography Award for Northern Canadian Research, Irene Woods Field Study Award, Laboratories without Borders and Undergrad-

uate Basic Mobility Award, and by the Polar Continental Shelf Program (PCSP). We are very thankful to Dr. Wayne Pollard, Dr. Christopher Omlon, Dr. Laura Thomson, Dr. Tim Moore, Mike Dalva, Alouette Zhang, Theresa Gossmann, Martine Dolmiere and the people of PCSP for their contributions to this project.

References

1. Müller F, Harrison J. Expedition Area - Axel Heiberg Island, Canadian Arctic Archipelago 1965;1:100000
2. Rudberg S. Distribution of small-scale periglacial and glacial geomorphological features on Axel Heiberg Island, Northwest Territories, Canada 1969
3. Andersen DT, Pollard WH, McKay CP, Heldmann J. Cold springs in permafrost on Earth and Mars. *J Geophys Res-Planet.* 2002;107(E3)
4. French HM. The periglacial environment. Hoboken, NJ: John Wiley & Sons, Inc.; 2018. Available from: <http://proxy.ohiolink.edu:9099/login?url=https://onlinelibrary.wiley.com/doi/book/10.1002/9781119132820>
5. Collins B, Wein G, Philippi T. Effects of Disturbance Intensity and Frequency on Early Old-Field Succession. *Journal of Vegetation Science.* 2001;12(5):721-8.
6. Kroetsch D, Wang C. Particle Size Distribution. Soil sampling and methods of analysis. Boca Raton, FL: Canadian Society of Soil Science. 2008; 713-26.
7. Kevan PG, Forbes BC, Kevan SM, Behan-Pelletier V. Vehicle Tracks on High Arctic Tundra: Their Effects on the Soil, Vegetation, and Soil Arthropods. *Journal of Applied Ecology.* 1995;32(3):655-67.
8. Frawley G. De Havilland Canada DHC-6 Twin Otter 2001; Available from: <https://www.airliners.net/aircraft-data/de-havilland-canada-dhc-6-twin-otter/181>
9. Beschel R, editor Observations on the time factor in interactions of permafrost and vegetation. Proceedings of the 1st Canadian conference on permafrost, Ottawa, Ontario Edited by JE Brown National Research Council of Canada, Ottawa, Ont; 1963.
10. Forbes B. Cumulative impacts of vehicle traffic on high arctic tundra: soil temperature, plant biomass, species richness and mineral nutrition. 1998;57:269-74.
11. Faure A. A new conception of the plastic and liquid limits of clay. *Soil & Tillage Research.* 1980;1:97-105.
12. Assouline S, Tessier D, Tavares-Filho J. Effect of compaction on soil physical and hydraulic properties: Experimental results and modeling. *Soil Science Society of America Journal.* 1997;61(2):390-8.
13. Lowery B, Schuler R. Temporal effects of subsoil compaction on soil strength and plant growth. *Soil Science Society of America Journal.* 1991;55(1):216-23.
14. Paquette M. Le rôle des formes périglaciaires dans l'hydrologie et l'évolution des pentes d'un désert polaire dans le Haut-Arctique canadien [Dissertation]. Montréal: Université de Montréal; 2018.
15. Warkentin B. Effects of compaction on content and transmission of water in soils. *Compaction of Agricultural Soils.* 1971.
16. Willis W, Raney W. Effects of compaction on content and transmission of heat in soils. *Compaction of Agricultural Soils.* 1971.
17. Shilts WW. Nature and genesis of mudboils, central Keewatin, Canada. *Canadian Journal of Earth Sciences.* 1978;15(7):1053-68.
18. Thermal Conductivity of Soil - Thermal Properties Application Available from: <https://thermtest.com/applications/soil-thermal-conductivity-tls>
19. Levy JS, Fountain AG, Gooseff MN, Welch KA, Lyons WB. Water tracks and permafrost in Taylor Valley, Antarctica: Extensive and shallow groundwater connectivity in a cold desert ecosystem. *Geological Society of America Bulletin.* 2011;123(11-12):2295-311.
20. Bintanja R. The impact of Arctic warming on increased rainfall. *Sci Rep-Uk.* 2018;8.
21. Becker MS, Pollard WH. Sixty-year legacy of human impacts on a high Arctic ecosystem. *Journal of Applied Ecology.* 2016;53(3):876-84.

Review Article

¹Guy's, King's, and St. Thomas'
School of Medical Education,
London, England

Keywords

Prosthetic limbs, powered prostheses, machine-body interface

Email Correspondence

nipun.mundkur@kcl.ac.uk

Nipun Mundkur¹

The Bionic Human: A Review of Interface Modalities for Externally Powered Prosthetic Limbs

Abstract

Background: The loss of a limb is a debilitating incident and can leave patients significantly disabled and often unable to perform activities of daily living. Prosthetic limbs can provide some modicum of normalcy back to their lives, and there has been much research over the past few decades into restoration of biomedical and physiological function with the use of externally powered and robotic prostheses. This review aims to explore the various approaches to machine-body interfacing that can be employed to achieve intuitive and meaningful control of these complex devices, and to discuss the individual benefits and drawbacks of each method.

Methods: Studies looked at include both primary and secondary sources of research. Identification was via a PubMed search for the terms “prosthetic limb”, “powered prostheses”, “myoelectric prostheses”, “neural interface”, “prosthetic somatosensory feedback”, and “brain-machine interface”, which resulted in a total of 3892 papers retrieved. Of these, 28 were retained as sources for this review. Selection was based on relevance to control of powered prostheses.

Summary: Significant strides have been made in expanding the choice of interface sites for bionic prosthesis control. Muscles, nerves, and the brain are all options, each with varying degrees of invasiveness and corresponding resolution of information obtained, and non-muscle interfacing prostheses may soon be commercially available. These advances have allowed for increasingly precise control of prosthetic limbs. However, this is limited by the challenge of returning sensory information from the prosthesis back to the user.

Introduction

Prosthetic limbs are artificial constructs that are attached in some way to the body of the user with the purpose of restoring at least some function of a lost or nonfunctional appendage. A bionic prosthesis is one that uses electrical signals from the body to move parts of the powered artificial limb. The most prevalent application of modern prosthetic limbs is the rehabilitation of injured soldiers who have lost limbs in the line of duty. A tragically common event, between 2003 and 2014, over 260 British soldiers in Afghanistan had to have amputations, with most having multiple limbs removed (1). However, amputation is not the only mechanism of functional limb loss. Traumatic spinal cord injuries and certain degenerative diseases often leave patients unable to move or sense their limbs. The ideal bionic limb would be able to reproduce all the functions of the limb that that these patients have lost. This includes all of the degrees of freedom of movement and somatosensory feedback to the user in all the modalities provided by a physiological limb. Furthermore, the prosthesis needs to be able to operate in all of the conditions a normal limb would, such as extremes of temperature and rain (2). The term “degree of freedom” in the context of this review is a plane of movement. For example, the upper limb, excluding the fingers, has a total of 7 degrees of freedom: 3 for the shoulder, 1 for the elbow, and 3 for the wrist.

Modern Prostheses and the Body-Prosthesis Interface

The key component of all prosthetic limbs, as well as all prostheses in general, is the interface between the user's body and the prosthesis. It is at this point that information can be exchanged between the user and the prosthesis, allowing purposeful movement of the limb and sensory feedback to the user. Classification is based on the site of interface. Body-powered prostheses are the oldest type and not considered bionic: they have existed for centuries and are simple pulley mechanisms where flexion and extension of an existing joint cause corresponding movement in the prosthetic distal joint. Myoelectric interfaces use signals from the remnant muscles

to control the prosthesis. Peripheral nerve interfaces utilise signals directly from nerves within the limb. Lastly, central nervous system (CNS) interfaces derive their information directly from the brain motor cortex. The amount of information that can be obtained (and therefore degree of control) is often dependent on the level of amputation. For example, it would be difficult to allow control of individual fingers in upper limb prostheses if the limb has been lost at the shoulder. However, if only the hand or distal forearm has been lost, fine control of fingers may be achieved, since many of the muscles and nerves that control the fingers in a normal limb are still present (2).

Another crucial aspect of the use of a prosthesis is the mechanism of attachment to the body of the user. In the past, the only way to achieve this was to use a harness that held the prosthesis against the limb stump. While harnesses are still widely used, particularly amongst users of passive prostheses, there are other options available today. Modern attachment mechanisms include suction sockets, elastic suspension sleeves, and osseointegration of the prosthesis. A suction socket is made of silicone and grips the remaining limb stump, while also being attached to the prosthesis (3). Elastic sleeves were developed as an alternative to suction sockets. They grip the stump due to their elastic property and are used to attach more lightweight prostheses. Osseointegration is the newest and most complex attachment mechanism, involving integration of the prosthesis with the remaining bone of the limb. This is done by drilling into the bone and screwing in the housing for the prosthesis. Unlike other methods of attaching the prosthesis, this is an invasive method involving implantation of foreign material into the patient's body. As such, there is a far greater risk of infection and other adverse reactions to the implanted part of the prosthesis, as well as the risk of damage occurring to adjacent structures during the procedure. While studies indicate low rates of infection among patients who have had the osseointegration procedure, the risk is still higher than with noninvasive attachment techniques (4). Nevertheless, osseointegration has several advantages, the first of which is that it is the most naturalistic of all the methods and reduces the discomfort experienced when prostheses are used for long periods of time. Furthermore, as this process allows a perma-

nent fixture to be available for prostheses to attach to, it can also provide a permanent site for interfacing the body and prosthesis. An example of this could be a peripheral nerve interface to nerves near the integration site, allowing both control and somatosensory feedback (5).

Myoelectric Interfaces

Myoelectric prostheses work by using information from remaining muscle to imitate their function and exercise control over the prosthetic. The information gathered is in the form of the electromyographic signals, which are caused by the generation of electrical potential within skeletal muscle cells during contraction. This information is gathered by surface electrodes, which then activate corresponding motors within the prosthesis that initiate movement, such as flexion/extension of the elbow and wrist and rotation of the latter (6). Currently there is only one pair of sensors in commercially available myoelectric prostheses, but multichannel sensors are being investigated to allow greater control (7). However, the activation of the motors is not homologous to muscle groups that perform the corresponding movement in a normal arm. As such, the information from the electrodes must be processed using certain algorithms, which attempts to recognize the movement that the user is trying to perform. However, this system is not entirely intuitive and therefore must be practiced and learned. This was the earliest bionic prosthesis type developed, and this was accomplished by scientists in the USSR in 1958 (8). As such, it is the only interface used in commercially available bionic prostheses; all other types are still experimental.

However, it is often found that the muscles that remain after loss of a limb are insufficient to allow precise control of a prosthesis. For example, if an arm is lost at the level of the shoulder, only a few upper limb muscles like the deltoid and pectorals will be available as sites for EMG detection, and this is not sufficient to allow utilisation of prostheses to perform complex movements or fine motor functions. A solution to this problem is Targeted Muscle Reinnervation (TMR). In TMR, the remaining nerves of the lost limb are surgically sited to separate locations within the residual muscles, with different nerves reinnervating different parts of the muscle. The signal from these nerves can then be picked up from the specific area of muscle that they reinnervate. The original nerves supplying this muscle are often removed if the action of the muscle is no longer needed. This procedure allows for natural amplification of the EMG signal, which allows for control of high degree-of-freedom prostheses without using invasive interfacing techniques and incurring the complications they are associated with. Furthermore, because the nerves that are used during operation of the prosthesis can be translocated, this allows for the patient to make movements of the prosthesis via more natural physiological pathways. This means that the patient can try to activate muscles that would normally be used for a movement, and because the nerves supplying these muscles have been rerouted and usable for the prosthesis, the patient feels a more intuitive control over the prosthesis, rather than using muscles normally unrelated to the intended movement. This results in faster and more precise control of the prosthesis (9).

Another issue with the use of myoelectric prostheses is that detection of EMG signals can be challenging and unreliable. This is due to anatomically close muscles whose contraction can cause interference in the signals, or the thickened scar tissue that is commonly found at sites of amputation (2). Naturally, this leads to imprecise and difficult control of the prosthesis. A potential solution to this is the Implantable Myoelectric Sensor (IMES), where electrodes are implanted within the muscles, allowing for considerably more sites to gather information from. Furthermore, due to their close association with specific motor units, interference from other muscles is greatly reduced. They are wireless and are powered by telemetry coils placed on the user. The sensors relay information to an external receiver that then controls motors of the prosthesis. They may be especially effective in prosthesis control when combined with TMR (7).

Peripheral Nerve Interfaces

Moving one level higher up the chain of motor control leads us to the peripheral nerves that supply the muscles of movement. Logically, interfacing with these structures would be a reasonable alternative to using EMG

signals, mimicking the way which normal limbs achieve movement. This is the principle behind the peripheral nerve interface. Furthermore, unlike a myoelectric interface, bidirectional information transfer is possible here, which opens up the prospect of somatosensory feedback. The peripheral nerve interface consists of a set of electrodes that are attached in some way to the remnants of the nerves that innervated the muscles of the limb, which, despite amputation, are still functional and follow the same pathways as before. The electrodes detect electroneurographic (ENG) signals from the nerves, which are then relayed to a processor that amplifies and modulates the information, then activates motors that control the movement of the limb (10).

However, there is no one way to go about obtaining information from peripheral nerves. The type of electrode used directly affects its selectivity and activation requirement. The main factor differentiating the electrode types is the degree of invasiveness into the nerve the electrodes are attached to. The first, and least invasive, is the cuff electrode. As the name implies, this type wraps around the entire circumference of the nerve and makes electrical contact with the epineurium, which is the outer sheath of the nerve. The electrode contacts do not enter the nerve, merely resting on its surface (11). The limitation of this type is that its circumferential nature causes it to have a low surface area, and hence lesser potential for detailed information gathering and selection of specific nerve fascicles. A potential solution to this is to flatten the nerve and apply a flat interface nerve electrode (FINE). This method takes advantage of the fact that nerves can be reshaped by constant forces on them over periods of time, and uses this to expand the surface area available and access more fascicles without increasing the invasiveness of the electrode (12).

Other peripheral nerve electrode types, such as Longitudinal Intrafascicular Electrodes (LIFE), are more invasive, with information being gathered from inside the nerve itself. To be specific, these electrodes are in contact with the interior of the nerve fascicles themselves. This increases the selectivity of the signals detected and reduces interference from other fascicles that may not transmit information that is useful for prosthesis control. The drawback of this is the danger of damaging the nerve during the implantation, which only increases with the invasiveness of the electrode (13). Due to its function as a nerve sheath, the epineurium is naturally insulating, meaning electrical signals are significantly harder to detect when the electrodes are outside the nerve rather than inside. For similar reasons, intrafascicular electrodes must penetrate the perineurium (which surrounds individual fascicles) as well. Due to the reduced resistance to electrical conduction within the fascicles as well as the increased number of electrode sites, even more information can be obtained compared to non-invasive (11). However, attempts to obtain more specific information have been made. An example of this is the sieve electrode, which is able to obtain information from each individual axon within a given nerve. However, the method of implantation is extremely invasive. It involves severing the entirety of the nerve, placing the sieve electrode between the cut ends and then letting the nerve regenerate with the electrode within it. The axons regenerate between the openings in the electrode, allowing discrete information to be obtained from each of them (14).

Cortical Interfaces

The final interface to be discussed is the cortical interface. In this modality, electrical signals for prosthesis control are obtained directly from the motor cortex of the brain itself, rather than peripheral tissues such as muscles or nerves. The advantage of using the central nervous system is that it provides an alternative site of information acquisition in patients with conditions that result in them being unable to send signals to said peripheral structures. Devices that can be controlled using cortical interfaces would be more useful for rehabilitation and quality-of-life improvement of patients affected by these conditions (15).

Cortical interfaces have sub-modalities of varying invasiveness. In most clinical scenarios where brain activity needs to be monitored, electroencephalographic (EEG) information, obtained via cutaneous electrodes, is used. However, while this method is noninvasive, the information gathered is of too low resolution to allow precise control of a prosthetic limb as the electrodes are excessively distant from the brain. A more invasive

modality is electrocorticographic (ECoG) information, where the information is obtained from platinum electrodes placed on the surface of the brain cortex. ECoG signals are currently the most widely used data acquisition modality for cortical interfaces. Lastly the most invasive modality is the low-field potential (LFP). Data is obtained via microelectrodes that pierce the cortex. These latter two modalities can be used in parallel as the signals they detect may encode different information about movements (16).

Motor control of the body arises from the part of the brain known as the motor cortex, which is located on the pre-central gyrus. Different parts of this cortex give signals to different parts of the body, and a relatively large cortical area is dedicated to control of the upper limb, the hand in particular, which is representative of the high dexterity of this appendage. The concept of these different cortical areas controlling different body parts can be used to develop a “map” (often represented as a person, the motor homunculus) for cortical brain-machine interface. This would mean that electrodes placed at a point that has been determined to control elbow movements, such as flexion and extension, would pick up signals when the patients wanted to perform said motion. These signals can then be processed into electronic commands for a prosthetic limb. This has led to the development of ECoG grids, which can be chronically implanted into the patient for years and overlaid onto the parts of their motor cortex that controls the limb to be replaced (17). However, the primary issue with the use of brain-machine interfaces is that electrode implantation causes reactive gliosis at the site. Gliosis is a form of tissue scarring where hyperplasia and hypertrophy of glial cells of the brain occur. This process is mediated by microglial cells, which have an immune-like function in the brain, and attach to the surface of the electrode after implantation in an attempt to destroy it. The result of this gliosis is the formation of a glial scar around the contact surface of the implanted electrodes, which interferes with and ultimately prevents the recording of cortical signals. In fact, although there is variation amongst patients, it has been found that around half of chronically implanted cortical electrodes are incapable of recording after several months, which results in reduction in the resolution of the information gathered and reduced control of the prosthesis over time (18).

Nonetheless, there are definite advantages associated with the use of cortical interfaces to control prostheses. Aside from the fact that it is the only type usable by those patients who cannot send peripheral signals, using cortical interfaces to control prostheses allows the patient to perform movements using the limb by merely “thinking” of doing so. This reduces the need for extensive training as the patients do not need to learn new and often unintuitive control schemes. Indeed, studies have shown that patients implanted with these interfaces were able to achieve reasonable control over a prosthesis with only an hour of practice, and the only aspect of control that had to be significantly changed were the algorithms that translated the patient’s ECoG signals into movement of the limb (19).

Somatosensory Feedback Systems

Somatosensory feedback signaling is necessary for the precise and coordinated movements observed in physiological limbs. These signals comprise large amounts of information, including the sensations of touch and pressure, the spatial position of limbs and joints, and temperature of the environment of the limb. As a whole, somatosensation of any part of the body allows the person to “embody” the appendage, and consider it part of their “self” (20).

As it currently stands, robotic limbs have almost-biomimetic movement and degrees of freedom; the issue is that patients are unable to utilize this to perform tasks requiring dexterity and precision (21). To achieve this, the prosthetic limb should have some form of somatosensory feedback, which would result in a significant increase in the user’s ability to perform complex non-preprogrammed motions (22). This is because the brain will be able to use the somatosensory feedback to modulate the signals it sends to move the prosthesis in a way that is far more intuitive and natural than the user having to observe their own movement and having to consciously decide when to interrupt or change the action. However, this retrograde transfer of information has proven to be considerably more challenging than in the usual human-to-prosthesis direction (23).

Much like achieving motor control of a prosthesis, there are multiple methods to deliver somatosensory information from prosthesis to body. In this situation communication must be with the nervous system as information transfer in myoelectric interfaces is unidirectional while peripheral nerves contain both afferent sensory axons and efferent motor axons, allowing for bidirectional information transfer. The peripheral sensory interface type can potentially be used in conjunction with osseointegration of the prosthetic limb house, which can provide a site for a permanent bidirectional interface between the user and the prosthesis (4). Lastly, brain-machine interfaces can also be used to transmit somatosensory information. However, while cortical interfaces for motor control made use of the motor cortex, such an interface for somatosensation must use the sensory cortex, which is located on the post-central gyrus of the brain.

As before with peripheral nerve interfaces, there are varying sub-types of electrodes, with their selectivity in accessing particular nerve fascicles increasing with their degree of invasiveness. However, there are ways to increase the selectivity of axon activation while minimizing invasiveness. For example, due to the decay of the electrical signal sent from a non-penetrative spiral electrode and the particular slew rate for the activation of a given axon, using electrical pulses of differing waveforms will allow axon types to be activated selectively depending on their size, myelination and distance from the electrode (24). In normal nerve physiology, the frequency of action potentials indicates the intensity of the stimuli being sensed, and this interface mimics that by modulating the frequency of the electrical signals delivered with the intensity detected. These methods have been shown to be able to elicit significantly localised sensations of touch of the phantom limb in amputees, allowing for improved motor control of prosthetic limbs (25).

Alternatively, if the somatosensory interface were to be of the cortical type, it would have to lie on the primary sensory cortex. This part of the cortex is further divided into four areas called Brodmann’s areas, to which neurons for particular sensations are localised. For example, one area is responsible for proprioception while another deals with pressure and light touch. Due to the highly precise and often complex nature of movements and sensations of the upper limb (particularly the hand), there is a considerable surface area of the sensory cortex devoted to processing of stimuli from this limb. With this knowledge, and the concept of the sensory homunculus (which is essentially the same idea as the motor homunculus described earlier), it would be possible to elicit localised sensations of various stimuli to parts of the limb that is to be replaced. This would be accomplished by using intracortical microstimulation (ICMS) to selectively activate neurons that are involved in the sensation of the stimuli to be delivered (26). As with peripheral nerve stimulation, the frequency of these electrical pulses indicate the intensity of the stimuli detected. However, in the cortical interface, modulation of the amplitude of the ICMS will also increase the intensity of the stimuli, and this is because a stronger signal will activate a larger number of neurons in proximity to the electrode (27).

The main difficulty in implementing cortical somatosensory interfaces is cortical plasticity. Due to cessation of afferent input from the lost limb, parts of the brain that normally process information from that limb begin doing so for other parts of the body instead. This means that electrical stimuli delivered to create sensations in part of the prosthesis may cause these sensations to be perceived as coming from elsewhere. Nevertheless, it has been shown that despite some changes in the organisation of the sensory cortex after limb loss, stimuli delivered to the appropriate areas will still cause sensations in the phantom limb (27).

Conclusion

For much of history, loss of a limb has been seen as a significant disability. Loss of legs led to negligible mobility and independence while loss of arms presented an incredible challenge to basic function. With the development of bionic prostheses, society stands at a point where the loss of a limb may soon become a far less debilitating incident, where much of the functionality of a normal limb can be restored.

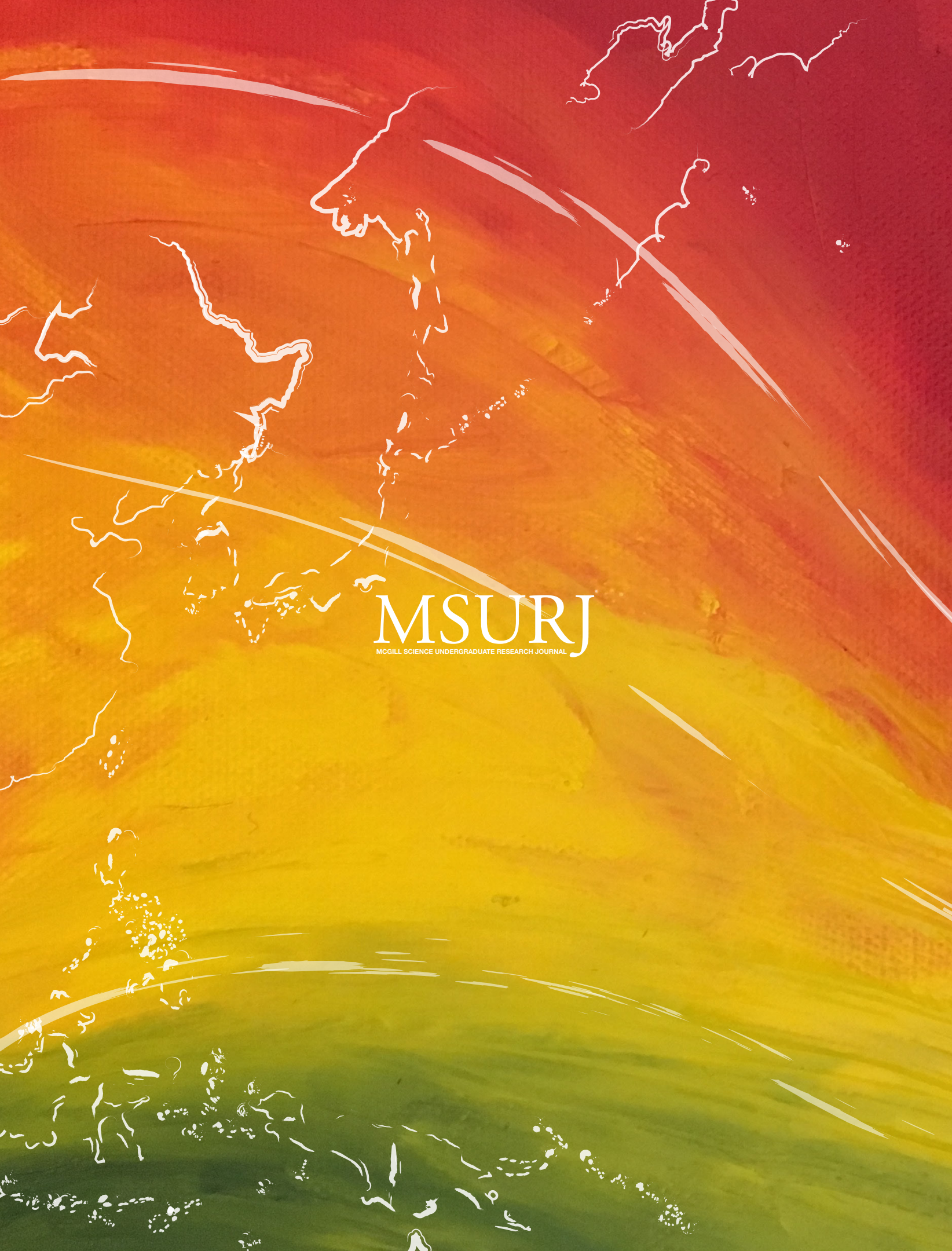
Looking to the future, it is likely that improvements in both the interfaces and the mechanical components of prostheses will lead to increasingly

biomimetic functioning of prosthetic limbs. However, it is increasingly possible that bionic limbs may not just replicate the normal functions of a limb, but augment it; functions that are not capable of being performed by normal limbs could be performed by bionic limbs: for example, the ability to lift heavier weights or manipulate the limb in novel manners may be introduced. Examples of such novel non-biomimetic functions can even be seen in prostheses today: users of the Bebionic myoelectric arm are able to rotate their wrists a complete 360° (28), and due to the materials used in modern prostheses, most allow users to grasp items such as hot or sharp objects that would normally injure a biological hand. Given that much of the funding for research in this field comes from the militaries of various countries, the potential for limb augmentation grows year by year.

27. Bensmaia S.J. Biological and bionic hands: natural neural coding and artificial perception. *Philos Trans R Soc Lond B Biol Sci.* 2014; 370(1677). DOI: 10.1098/rstb.2014.0209
28. Belter J.T. et al. Mechanical design and performance specifications of anthropomorphic prosthetic hands: A review. *J Rehabil Res Dev.* 2013; 50(5): 599-618.

References

1. Edwards R.D. et al. What Is the Magnitude and Long-term Economic Cost of Care of the British Military Afghanistan Amputee Cohort? *Clin Orthop Relat Res.* 2015; 473(9): 2848-55.
2. Hutchinson D.T. The Quest for the Bionic Arm. *J Am Acad Orthop Surg.* 2014; 22(6): 346-51.
3. Fillauer C.E., Pritham C.H., Fillauer K.D. Evolution and Development of the Silicone Suction Socket (3S) for Below-Knee Prostheses. *J Prosthet Orthot.* 1989; 1(2): 92-103.
4. Brånemark R. et al. A novel osseointegrated percutaneous prosthetic system for the treatment of patients with transfemoral amputation: A prospective study of 51 patients. *Bone Joint J.* 2014; 96-B(1): 106-13.
5. Khemka A. et al. Osseointegrated total hip replacement connected to a lower limb prosthesis: A proof-of-concept study with three cases. *J Orthop Surg Res.* 2016; 11(13). DOI: 10.1186/s13018-016-0348-3
6. Englehart K., Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans Biomed Eng.* 2003; 50(7): 848-54.
7. McDonnall D. et al. Implantable multichannel wireless electromyography for prosthesis control. *Conf Proc IEEE Eng Med Biol Soc.* 2012; 24: 1350-3.
8. Hargrove L.J. et al. Toward the development of a neural interface for lower limb prosthesis control. *Conf Proc IEEE Eng Med Biol Soc.* 2009; 21: 2111-4.
9. Kuiken T.A. et al. Targeted Muscle Reinnervation for Real-Time Myoelectric Control of Multifunction Artificial Arms. *JAMA.* 2009; 301(6): 619-28.
10. del Valle J., Navarro X. Interfaces With the Peripheral Nerve for the Control of Neuroprostheses. *International Review of Neurobiology.* 2013; 109: 63-83.
11. Micera S., Carpaneto J., Raspopovic S. Control of hand prostheses using peripheral information. *IEEE Rev Biomed Eng.* 2010; 3: 48-68.
12. Tyler D.J., Durand D.M. Functionally selective peripheral nerve stimulation with a flat interface nerve electrode. *IEEE Trans Neural Syst Rehabil Eng.* 2002; 10(4): 294-303.
13. Micera S. et al. Decoding of grasping information from neural signals recorded using peripheral intrafascicular interfaces. *J Neuroeng Rehabil.* 2011; 8(53). DOI: 10.1186/1743-0003-8-53
14. Navarro X. et al. Stimulation and recording from regenerated peripheral nerves through polyimide sieve electrodes. *J Peripher Nerv Syst.* 1998; 3(2): 91-101.
15. Hochberg L.R. et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature.* 2012; 485(7398): 372-5.
16. Fifer M.S. et al. Toward electrocorticographic control of a dexterous upper limb prosthesis: Building brain-machine interfaces. *IEEE Pulse.* 2012; 3(1): 38-42.
17. Wang W. et al. An electrocorticographic brain interface in an individual with tetraplegia. *PLoS One.* 2013; 8(2). DOI: 10.1371/journal.pone.0055344
18. Leach J.B., Achyuta A.K.H., Murthy S.K. Bridging the Divide between Neuroprosthetic Design, Tissue Engineering and Neurobiology. *Front Neuroengineering.* 2009; 2(18). DOI: 10.3389/neuro.16.018.2009
19. Yanagisawa T. et al. Real-time control of a prosthetic hand using human electrocorticography signals. *J Neurosurg.* 2011; 114(6): 1715-22.
20. Proske U., Gandevia S.C. The Kinaesthetic Senses. *J Physiol.* 2009; 17(587): 4139-46.
21. Johannes M.S. et al. An overview of the developmental process for the Modular Prosthetic Limb. *JHU APL Technical Digest.* 2011; 30(3): 207-16.
22. Tabot G.A. et al. Restoring the sense of touch with a prosthetic hand through a brain interface. *Proc Natl Acad Sci USA.* 2013; 110(45): 18279-84.
23. Tan D.W. et al. A neural interface provides long-term stable natural touch perception. *Sci Trans Med.* 2014; 6(257): 257ra138.
24. Fang Z.P., Mortimer J.T. Selective activation of small motor axons by quasi-trapezoidal current pulses. *IEEE Trans Biomed Eng.* 1991; 38(2): 168-74.
25. Tyler D.J. Neural interfaces for somatosensory feedback: bringing life to a prosthesis. *Curr Opin Neurol.* 2015; 28(6): 574-81.
26. Johnson L.A. et al. Direct electrical stimulation of somatosensory cortex in humans using electrocorticography electrodes: a qualitative and quantitative report. *J Neural Eng.* 2013; 10(3). DOI: 10.1088/1741-2560/10/3/036021



MSURJ

MCGILL SCIENCE UNDERGRADUATE RESEARCH JOURNAL