Microbial dynamics in a High-Arctic glacier forefield: a combined field, laboratory, and modelling approach.

James A. Bradley 1,2, Sandra Arndt 2, Marie Šabacká 1, Liane G. Benning 3,4, Gary L. Barker 5, Joshua J. Blacker 3, Marian L. Yallop 5, Katherine E. Wright 1, Christopher M. Bellas 1, Jonathan Telling 1, Martyn Tranter 1, Alexandre M. Anesio 1

1 Bristol Glaciology Centre, School of Geographical Sciences, University of Bristol, BS8 1SS, UK
2 BRIDGE, School of Geographical Sciences, University of Bristol, BS8 1SS, UK
3 School of Earth and Environment, University of Leeds, LS2 9JT, UK
4 GFZ, German Research Centre for Geosciences, 14473 Potsdam, Germany
5 School of Biological Sciences, University of Bristol, BS8 1SS, UK

Corresponding author: James A. Bradley, email: j.bradley@bristol.ac.uk

Abstract: Modelling the development of soils in glacier forefields is necessary in order to assess how microbial and geochemical processes interact and shape soil development in response to glacier retreat. Furthermore, such models can help us predict microbial growth and the fate of Arctic soils in an increasingly ice-free future. Here, for the first time, we combined field sampling with laboratory analyses and numerical modelling to investigate microbial community dynamics in oligotrophic proglacial soils in Svalbard. We measured low bacterial growth rates and growth efficiencies (relative to estimates from Alpine glacier forefields), and high sensitivity to soil temperature (relative to temperate soils). We used these laboratory measurements to inform parameter values in a new numerical model and significantly refined predictions of microbial and biogeochemical dynamics of soil development over a period of roughly 120 years. The model predicted the observed accumulation of autotrophic and heterotrophic biomass. Genomic data indicated that initial microbial communities were dominated by bacteria derived from the glacial environment, whereas older soils hosted a mixed community of autotrophic and heterotrophic bacteria. This finding was simulated by the numerical model, which showed that active microbial communities play key roles in fixing and recycling carbon and nutrients. We also demonstrated the role of allochthonous carbon and microbial necromass in sustaining a pool of organic material, despite high heterotrophic activity in older soils. This combined field, laboratory and modelling approach demonstrates the value of integrated model-data studies to understand and quantify the functioning of the microbial community in an emerging High-Arctic soil ecosystem.

Key words
Glacier forefield
Microbial dynamics
Soil development
Numerical modelling
Integrated field-laboratory-modelling
SHIMMER
1. Introduction

Polar regions are particularly sensitive to anthropogenic climate change (Lee, 2014) and have experienced accelerated warming in recent decades (Johannessen et al., 2004; Serreze et al., 2000; Moritz et al., 2002). The response of terrestrial Polar ecosystems to this warming is complex, and research to understand the response of terrestrial ecosystems in high latitudes to environmental change is of increasing importance. A visible consequence of Arctic warming is the large-scale retreat of glacier and ice cover (ACIA, 2005; Paul et al., 2011; Staines et al., 2014; Dyurgerov and Meier, 2000). From underneath the ice, a new terrestrial biosphere emerges, playing host to an ecosystem which may exert an important influence on biogeochemical cycles, and more specifically atmospheric CO₂ concentrations and associated climate feedbacks (Dessert et al., 2003; Anderson et al., 2000; Smittenberg et al., 2012; Berner et al., 1983).

Numerous studies have attempted to characterize the physical and biological development of recently exposed soils using a chronosequence approach, whereby a transect perpendicular to the retreating ice snout represents a time sequence with older soils at increasing distance from the ice snout (Schulz et al., 2013). We have recently shown that microbial biomass and macronutrients (such as carbon, phosphorus and nitrogen) can accumulate in soils over timescales of decades to centuries (Bradley et al., 2014). In such pristine glacial forefield soils the activity of microbial communities is thought to be responsible for this initial accumulation of carbon and nutrients. Such an accumulation facilitates colonization by higher order plants, leading to the accumulation of substantial amounts of organic carbon (Insam and Haselwandter, 1989). However, organic carbon may also be derived from allochthonous sources such as material deposited on the soil surface (from wind, hydrology, precipitation and ornithogenic sources) and ancient organic pools derived from under the glacier (Schulz et al., 2013). Nevertheless, the relative significance of allochthonous and autochthonous sources of carbon to forefield soils, as well as their effect on ecosystem behaviour are so far still poorly understood (Bradley et al., 2014). Moreover, cycling of bioavailable nitrogen (which is derived from active nitrogen-fixing organisms, allochthonous deposition, and degradation of organic substrates) and phosphorus (liberated from the weathering of minerals and decomposition of organic substrates) are similarly poorly quantified.

Several studies have observed shifts in the microbial community inhabiting pro-glacial soils of various ages (Zumsteg et al., 2012; Zumsteg et al., 2011). This was expressed in increasing rates of autotrophic and bacterial production with soil age (Schmidt et al., 2008; Zumsteg et al., 2013; Esperschutz et al., 2011; Frey et al., 2013) and the overall decline in quality of organic substrates in older soils (Goransson et al., 2011; Insam and Haselwandter, 1989). However, current evidence is limited to mostly descriptive approaches, which may be challenging to interpret due to inherent difficulties in disentangling interacting microbial and geochemical processes across various temporal and spatial scales. Furthermore, the inherent heterogeneity of glacial forefield soils makes the development of a single conceptual model that fits all challenging. Accordingly, pro-glacial biogeochemical processes that dominate such systems
remain poorly quantified and highly under-explored. This current lack of understanding limits our ability

to predict the future evolution of these emerging landscapes and the potential consequences on global
climate. Numerical models present an opportunity to expand our knowledge of glacier forefield
ecosystems by analytically testing the hypotheses that arise from observations, extrapolating,
interpolating and budgeting processes, rates and other features to explore beyond the possibility of
empirical observation (Bradley et al., 2016). With such a model we can then also explore the sensitivity
and resilience of these ecosystems to environmental change.

Here, we have combined field observations, with laboratory incubations and elemental measurements
as well as genomic analyses and used these in a numerical model to investigate the development of
soils in a glacial forefield. The present study forms an important part of the integrated and iterative
model-data approach outlined in the model description and testing (Bradley et al, 2015) whereby initial
model development was informed by decades of empirical research, new data and laboratory
experiments (presented here) are used to refine and inform model simulations, and so forth. With this
data we refined some model parameters in the recently developed Soil biogeochemical Model for
Microbial Ecosystem Response (SHIMMER 1.0; Bradley et al. (2015)) model and applied this to the
emerging forefield of the Midtre Lovénbreen glacier in Svalbard. Pioneer soils in the High-Arctic and
Antarctica, such as the Midtre Lovénbreen forefield, are ideal sites to test this field-laboratory-model
approach due to the lack of vegetation during initial stages of soil development, as the presence of
vegetation would obscure the microbial community dynamics and considerably alter the physical
properties of the soil (Brown and Jumpponen, 2014; Ensign et al., 2006; King et al., 2008; Kastovska
et al., 2005; Schutte et al., 2009; Duc et al., 2009). The model development was informed by decades
of empirical research on glacier forefield soils, and has already been tested and validated using
published datasets from the Damma Glacier in Switzerland and the Athabasca Glacier in Canada. A
thorough sensitivity analysis highlighted the most important parameters to constrain in order to make
further predictions more robust. All our model parameter values are specific to individual, local model
conditions and inherently contain necessary model simplifications, abstractions and assumptions.
Nevertheless, our earlier sensitivity analyses revealed the following highly sensitive key parameters as
the most important to constrain through measurements: the maximum heterotrophic growth rate ($I_{\text{maxH}}$),
the bacterial growth efficiency (BGE, parameter $Y_{\text{H}}$) and the temperature response ($Q_{10}$).

Therefore, in this current study, we combined detailed field measurements with specifically designed
laboratory experiments and quantified values for these three parameters with a specific set of soils from
for the Midtre Lovénbreen forefield. The laboratory experiments and measurements were conducted
with the objective to better constrain these sensitive parameters. We then ran model simulations in
order to explore the ranges of model output and refine model predictions compared to the previous
range identified in Bradley et al (2015). Next, we examined model output to explore the microbial and
biogeochemical dynamics of recently exposed soils in the Midtre Lovénbreen catchment and evaluate
two main hypotheses. First, we tested the hypothesis that microbial biomass in recently exposed soils
accumulates due to in situ bacterial growth and activity. It is commonly observed in glacier forefields
that microbial biomass accumulates with increasing soil age following exposure (Bernasconi et al., 2011; Schulz et al., 2013; Bradley et al., 2014). This study provides a new quantitative and process-focused approach to examine in situ growth in pioneer ecosystems, and assess the role of different functional groups in biomass accumulation. Second, we tested the hypothesis that carbon fluxes in very recently exposed soils are low, and are dominated by (abiotic) deposition of allochthonous substrate, whereas carbon fluxes are high in older soils due to increased microbial (biotic) activity (such as microbial growth, respiration and cell death). Increased soil carbon fluxes with soil age have been linked to microbial activity from the forefield of the Damma Glacier, Switzerland (Smittenberg et al., 2012; Guelland et al., 2013b). With this combined model, field and lab study, we were able to estimate carbon fluxes between ecosystem components with daily resolution, and provide new insight into the interplay of processes that contribute to net ecosystem production and soil organic carbon stocks in a High-Arctic system.

2. Methods

2.1. Study site and sampling
Midtre Lovénbreen is an Arctic polythermal valley glacier on the south side of Kongsfjorden, Western Svalbard (latitude 78°55’N, longitude 12°10’E) (Fig. 1). The Midtre Lovénbreen catchment is roughly 5 km East of Ny-Ålesund, where several long-term monitoring programs have provided a wealth of contextual information. Midtre Lovénbreen has experienced negative mass balance throughout much of the 20th century. Since the end of the Little Ice Age (maximum in Svalbard in the 1900s) the deglaciated surface area of the Midtre Lovénbreen catchment has increased considerably in response to warming mean annual temperatures. This continues to the present day. Between 1966 and 1990 ~2.3 km² of land was exposed (Fleming et al., 1997; Moreau et al., 2008). We used a chronosequence approach to determine ages for soils based on satellite imagery (Landsat TM 7) and previously determined soil ages by aerial photography and carbon-14 dating techniques in Hodkinson et al. (2003). Soil samples were collected along a transect perpendicular to the glacier snout, representing soil ages of 0, 3, 5, 29, 50, and 113 years (Fig. 1) during the field season (18 July to 29 August 2013). At each of the 6 sites along the chronosequence, 10 meter traverses roughly parallel to the glacier snout were established and at each site 3 soil plots were sampled (using ethanol sterilized sampling equipment). After removing the > 2 cm rock pieces at each site, about 100 grams of soil was collected from the top 15 cm and immediately placed into sterile high-density polyethylene bags (Whirl-Pak (Lactun, Australia)) and into a cool box partially filled with cool packs and dry ice. Samples were immediately frozen and stored at -20°C on return to the UK Arctic Research station in Ny-Ålesund (no longer than 5 hours after collection). Samples were transported frozen on dry ice to the laboratories in the Universities of Bristol and Leeds (UK).

2.2. Laboratory analyses
For bacterial abundance, samples were thawed and aliquots (100 mg) were immediately transferred into sterile 1.5 mL micro-centrifuge (Eppendorf) tubes, where they were diluted with 900 μL of Milli-Q water (0.2 μm filtered) and immediately fixed in 100 μL glutaraldehyde (0.2 μm filtered, 2.5% final
concentration). Samples were then vortexed for 10 seconds and sonicated for 1 minute at 30°C to facilitate cell detachment from soil particles. Then 10 µL fluorochrome DAPI (4', 6-diamidino-2-phenylindole) was added to half of the samples, tubes were vortexed briefly (3 seconds) and incubated in the dark for 10 minutes, to be counted under UV light. The other half of each sample remained untreated, for counting under auto-fluorescent light for photosynthetic pigmentation. Samples were vortexed for 10 seconds and let stand for a further 30 seconds to ensure a well-mixed solution, prior to filtering 100 µL of the mixed liquid sample onto black Millipore Isopore membrane filters (0.2 µm pore size, 25 mm diameter), rinsed with a further 250 µL of Milli-Q water (0.2 µm filtered). Bacterial cells were then counted using an Olympus BX41 microscope at 1000 times magnification. The filtering apparatus was washed out with Milli-Q water between each filtration, and negative control samples, prepared using Milli-Q water, were included into each series. A negative control was a sample with no visible stained or auto-fluorescing cells. Thirty random grids (each 10^4 μm²) were counted per sample. Cell morphologies were measured and cell volume was estimated and converted to carbon content according to Bratbak and Dundas (1984) (see Supplementary Information). Separate aliquots of soil from each site were weighed after thawing and then dried at 105°C to obtain an estimate of soil moisture content.

Environmental DNA was isolated from at least 3 replicates for each soil age using MoBio PowerMax® Soil DNA Isolation Kit and by following the instruction manual. 5 to 10 g of soil was used per sample to isolate DNA. Following extraction, DNA was diluted to equimolar concentrations, so that all samples contained 7.0 ng of DNA, prior to amplification. The isolated genomic DNA was amplified with bacterial primers 515f (5'-GTGYCAGCMGCCGCGGTAA-3') and 926r (5'-CCGYCAATTYMTTTRAGTTT-3') (Caporaso et al., 2012), creating a single amplicon of ~400 bp. The reaction was carried out in 50 µL volumes containing 0.3 mg mL⁻¹ Bovine Serum Albumin, 250 µM dNTPs, 0.5 µM of each primer, 0.02 U Phusion High-Fidelity DNA Polymerase (Finnzymes OY, Espoo, Finland) and 5x Phusion HF Buffer containing 1.5 mM MgCl₂. The following PCR conditions were used: initial denaturation at 95°C for 5 minutes, followed by 25 cycles consisting of denaturation (95°C for 40 seconds), annealing (55°C for 2 minutes) and extension (72°C for 1 minute) and a final extension step at 72°C for 7 minutes. Following amplification, all amplicons were normalized to the same concentration (9.0 ng) before sequencing. Samples were sequenced using the Ion Torrent platform (using Ion 318v2 chip) at Bristol Genomics facility at the University of Bristol. Samples were barcoded in the PCR stage and demultiplexed in QIIME using split_libraries.py code (Caporaso et al., 2010). A non-barcoded library was prepared from the amplicon pool using Life technologies Short Amplicon Prep Ion Plus Fragment Library Kit. The template and sequencing kits used were: Ion PGM Template OT2 400 Kit and Ion PGM Sequencing 400 kit. The sequencing yielded 4.38 million reads. The 16S sequences were further processed using MOTHUR (v. 1.35) and QIIME pipelines (Schloss et al., 2009; Caporaso et al., 2010). Initially, low quality or too long and too short sequences were removed in MOTHUR. Chimeric sequences were identified and removed using UCCHIME (Edgar et al., 2011). QIIME was used to cluster reads into operational taxonomical units (OTUs) using the pick_closed_reference_otus.py command. The final OTU table contained between 2329-43269 reads for each sample. The varying sequence coverage among samples was adjusted by
normalizing the OTU matrix to the same total abundance of sequences for each sample (2329) using the command normalize_table.py in QIIME. Sequences were clustered into OTUs based on at least 97% sequence similarity, and assigned taxonomical identification against Greengenes bacterial database (McDonald et al., 2012). The result was a biom format file containing the taxonomic information for each OTU as well as OTU frequency per sample. Sequence data is available on NCBI under project ID PRJNA341831, and individual accession numbers are provided in the Supplementary Information.

The carbon contents in the year 0 soils were analysed with a Carlo-Erba elemental analyser (NC2500) at the German Research Center for Geosciences, Potsdam, Germany. The soils were oven dried at 40°C for 48 hours, sieved to <7 mm and crushed using a TEMA disk mill to achieve size fractions of < 20 μm. Total organic carbon (TOC) was analysed after reacting the powders with a 10% HCl solution for 12 hours to remove inorganic carbonates.

2.3. Determination of maximum growth rates

The microbial activity was determined in 113 year old soil samples after they were thawed (in the dark at 5°C to mimic typical field temperature) for 168 hours. This age was chosen because these soil samples were assumed to be the ones with the highest microbial biomass and activity and thus the most practical for all laboratory measurements. In order to mitigate the effect of variability derived from differences in soil properties between soil ages (that will later be predicted by the model), laboratory experiments were conducted on a single soil age, with replicate incubations to assess the possible variability in rates (and thus parameter values) that can be attributed to experimental procedures and measurement techniques.

Aliquots of the soils were divided into petri dishes (25 g of soil (wet weight) into each petri dish) for subsequent treatments. In order to alleviate nutrient limitations and measure maximum growth rates, four different nutrient conditions were simulated: (1) no addition of nutrients, (2) low (0.03 mg C g⁻¹, 0.008 mg N g⁻¹, 0.02 mg P g⁻¹), (3) medium (0.8 mg C g⁻¹, 0.015 mg N g⁻¹, 0.1 mg P g⁻¹) and (4) high additions (2.4 mg C g⁻¹, 0.024 mg N g⁻¹, 0.3 mg P g⁻¹). The ranges and concentrations were informed by similar experiments in recently exposed proglacial soils at the Damma Glacier, Switzerland (Goransson et al., 2011). Nutrients (C₆H₁₂O₆ for C, NH₄NO₃ for N and KH₂PO₄ for P) (Sigma, quality ≥99.0%) were dissolved in 2 mL Milli-Q water (0.2 µm filtered), and mixed into the soils using an ethanol-sterilized spatula. Samples were incubated in the dark for a further 72 hours with the lids on at 25°C, the reference temperature (T_ref) at which all rates are defined in SHIMMER prior to adjustment with the temperature dependency expression (using Q₁₀) (Bradley et al., 2015). In order to derive a value for I_max, we were obligated to perform growth incubations at T_ref (25°C) despite this being a more typical soil temperature of Alpine soils rather than High-Arctic soils (see Fig. S3 (c)). However, we are confident that by deriving a Q₁₀ value based on incubations of the same soils encapsulating typical (5°C) to high (25°C) soil temperatures, we can numerically derive appropriate actual growth rates from the maximum growth rate (at T_ref). We are confident that the major outcomes and conclusions of this study are not
affected by high incubation temperatures since measured growth rates at high temperatures are appropriately scaled using the $Q_{10}$ formulation as measured experimentally. Throughout the whole incubation time, at 24 hour intervals, additional 2 mL aliquots of Milli-Q water (0.2 µm filtered) were added to maintain approximate soil moisture conditions in each sample.

In these samples bacterial production was estimated by the incorporation of $^3$H-leucine using the microcentrifuge method detailed in Kirchman (2001). After the initial 72 hour incubation period quadruplicate sample aliquots from the petri dish incubations and two trichloroacetic acid (TCA) killed control samples were incubated for 3 hours at $T_{ref}$ (25°C) for every nutrient treatment. Approximately 50 mg of soil was transferred to sterile micro-centrifuge tubes (2.0 mL, Fischer Scientific). Milli-Q (0.2 µm pre-filtered) water and $^3$H-leucine was added to a final concentration of 100 nM (optimum leucine concentration was pre-determined by a saturation experiment, Fig. S1, Supplementary Information). The incubation was terminated by the addition of TCA to each tube. Tubes were then centrifuged at 15,000 g for 15 minutes, the supernatant was aspirated with a sterile pipette and removed, and 1 ml ice-cold 5% TCA was added to each tube. Tubes were then centrifuged again at 15,000g for 5 minutes, before again aspirating and removing the supernatant. 1ml ice-cold 80% ethanol was added and tubes were centrifuged at 15,000 g for 5 minutes, before the supernatant was aspirated and removed again and tubes were left to air dry for 12 hours. Finally, 1 mL of scintillation cocktail was added, samples were vortexed, and then counted by liquid scintillation (Perkin Elmer Liquid Scintillation Analyzer, Tri-Carb 2810 TR). Radiosotope activity of TCA-killed control samples was always less than 1.1% of the measured activity in live samples. There was a positive correlation between the amount of sediment added to the tubes and background counts representing disintegrations per minute (DPM). Counts were individually normalized by the amount of sediments (corrected for dry weight) used in each sample to discount for background DPM. Leucine incorporation rates were converted into bacterial carbon production following the methodology of Simon and Azam (1989). Bacterial abundance was estimated from each treatment after the 72 hour incubation period by microscopy. Five samples from each petri dish were counted for each nutrient treatment with negative controls yielding no detectable cells. One-way ANOVA (with post-hoc Tukey HSD) statistical tests were used for evaluations of the variability from the multiple treatments.

2.4. Temperature response

Microbial community respiration was determined by measuring CO$_2$ gas exchange rates in airtight incubation vials. Soil samples from the 113 year old site were defrosted and divided (25 g wet weight) in petri dishes as above, and 2 mL of Milli-Q water (0.2 µm filtered) was added (to maintain consistency of soil moisture with determination of bacterial production above). Samples were incubated at 5°C (T$_1$) and 25°C (T$_2$) in the dark for a further 72 hours. 2mL of 0.2 µm pre-filtered Milli-Q water was added to the T$_1$ sample (3 mL for T$_2$) at 24, 48 and 72 hours to maintain approximate soil moisture content. Two separate killed control tests (one furnaced at 450°C for 4 hours, and one autoclaved (3 cycles at 121°C)) were incubated at T$_1$ and T$_2$. Quintuple live and killed samples (roughly 1 g) were transferred into cleaned 20 mL glass vials (rinsed in 2% Decon, submersed in 10% HCl for 24 hours, rinsed 3 times...
with Milli-Q water and furnace at 450°C for 4 hours). These were sealed (9°C, atmospheric pressure, ambient CO₂ of 405 ppm) with pre-sterilized Belco butyl stoppers (pre-sterilized by boiling for 4 hours in 1M sodium hydroxide) and crimped shut with aluminium caps. Sealed vials were then incubated at T₁ and T₂ for 24 hours in darkness. After 24 hours, the headspace gas was removed with a gas-tight syringe and immediately analysed on an EGM4 gas analyser (PP Systems, calibrated using gas standards matching the expected range, precision 1.9%, 2*SE). Empty pre-sterilized vials were also incubated and analysed. Following gas analysis, vials were opened and dried to a constant weight at 105°C to estimate moisture content and thus dry soil weight of these aliquots. Headspace CO₂ change (ppm) was converted to microbial respiration using the ideal gas law (n=PV/RT), assuming negligible changes in soil pore water pH (and therefore CO₂ solubility) during the incubation. CO₂ headspace changes resulting from killed controls and blanks were < 70% of the changes resulting from the incubations at T₁, and >7% of the changes observed at T₂. One-way ANOVA (with post-hoc Tukey HSD) statistical tests were used for comparison of multiple treatments. No significant differences in CO₂ headspace change between killed controls at T₁ and T₂ were detected (P=0.95).

2.5. Microbial Model: SHIMMER

SHIMMER (Bradley et al., 2015) mechanistically describes and predicts transformations in carbon, nitrogen and phosphorus through aggregated components of the microbial community as a system of interlinked ordinary differential equations. The model contains pools of microbial biomass, organic matter and both dissolved inorganic and organic nitrogen and phosphorus (Table 1). It categorizes microbes into autotrophs (A₁-3) and heterotrophs (H₁-3), and further subdivides these based on 3 specific functional traits. Microbes derived from underneath the glacier (referred to as “glacial microbes”) are termed A₁ and H₁. A₁ are chemolithoautotrophic, obtaining energy from the oxidation and reduction of inorganic compounds and carbon from the fixation of carbon dioxide. In contrast, H₁ rely on the breakdown of organic molecules for energy to support growth. A₂ and H₂ represent autotrophic and heterotrophic microbes commonly found in glacier forefield soils with no “special” characteristics, and will be referred to as “soil microbes”. A₃ and H₃ are autotrophs and heterotrophs that are able to fix atmospheric N₂ gas as a source of nitrogen in cases when dissolved inorganic nitrogen (DIN) stocks become limiting. Available organic substrate is assumed to be derived naturally from dead organic matter and allochthonous inputs. Labile compounds are immediately available fresh and highly reactive material, rapidly turned over by the microorganisms (S₁, ON₁, OP₁). Refractory compounds are less bioavailable and represent the bulk of substrate present in the non-living organic component of soil (S₂, ON₂, OP₂). A conceptual diagram showing the components and transfers of SHIMMER is presented in the Supplementary Information (Fig. S2).

Microbial biomass responds dynamically to changing substrate and nutrient availability (expressed as Monod-kinetics), as well as changing environmental conditions (such as temperature and light). A Q₁⁰ temperature response function (T) is affixed to all metabolic processes including growth rates and death rates (Bradley et al., 2015), thus effectively slowing down or speeding up all life processes as temperature changes (Soetaert and Herman, 2009; Yoshitake et al., 2010; Schipper et al., 2014). Light
limitation is expressed as Monod kinetics. The following external forcings drive and regulate the system’s dynamics:

- Photosynthetically-active radiation (PAR) (wavelength of approximately 400 to 700 nm) (W m$^{-2}$).
- Snow depth (m).
- Soil temperature (°C).
- Allochthonous inputs (µg g$^{-1}$ day$^{-1}$).

The model is 0-D and represents the soil as a homogeneous mix. Thus, light, temperature, nutrients, organic compounds and microbial biomass are assumed to be evenly distributed.

Soil temperature (at 1 cm depth) for the entire of 2013 is provided by Alfred Wegener Institute for Polar and Marine Research (AWI) from the permafrost observatory near Ny-Ålesund, Svalbard. Similarly, PAR for 2013 are measured at the AWI surface radiation station near Ny-Ålesund, Svalbard. Averaged daily snow depth for 2009 to 2013 is provided by the Norwegian Meteorological Institute (eKlima). Allochthonous nutrient fluxes (inputs and leaching) are estimated based on an evaluation of nutrient budgets of the Midtre Lovénbreen catchment (Hodson et al., 2005) in which budgets for nutrient deposition rates and runoff concentrations are measured over two full summer-winter seasons and residual retention rates (excess of inputs) or depletion rates (excess of outputs) are inferred. The bioavailability of allochthonous material is assumed to be the same as initial material and microbial necromass.

Initial conditions were informed by analysis of 0-years-of-exposure soil collected adjacent to the ice snout, and initial values for all state variables are presented in Table 1. Initial microbial biomass was estimated by microscopy as described above. Initial community structure was derived by 16S analysis of year-0 soils. An initial value for carbon substrate ($S_1 + S_2$) was estimated based on the average TOC content of year-0 soil. Bioavailability of model TOC was assumed to be 30% labile ($S_1$) and 70% refractory ($S_2$) (for consistency with Bradley et al. (2015)). Organic nitrogen (ON) and organic phosphorus (OP) were assumed to be stoichiometrically linked by the measured C:N:P ratio from the Damma Glacier forefield (from which the model was initially developed and tested (Bradley et al., 2015)). An initial value for DIN was taken from a previous evaluation of Svalbard tundra nitrogen dynamics, whereby the lowest value is taken to represent the soil of least development, according to traditional understanding of glacier forefields (Alves et al., 2013; Bradley et al., 2014). An initial value for dissolved inorganic phosphorous (DIP) was established stoichiometrically from previous model development and testing.

Model implementation and set-up is described in more detail in the Supplementary Information.

2.6. Model parameters
Maximum heterotrophic growth rate \( I_{\text{maxH}} \) (day\(^{-1}\)) was estimated by scaling the measured rate of bacterial production (\( \mu g \text{ C g}^{-1} \text{ day}^{-1} \)) (converted to dry weight) with total heterotrophic biomass (\( \mu g \text{ C g}^{-1} \)). Nutrient addition alleviates growth limitations as defined in SHIMMER (Bradley et al., 2015); thus bacterial communities can be assumed to be growing at \( I_{\text{maxH}} \) under experimental conditions.

\( Y_H \) represents heterotrophic BGE, and was estimated according to the equation:

\[
Y_H = \frac{BP}{BP + BR}
\]

(1)

Where \( BP \) is and \( BR \) are measured bacterial production and measured bacterial respiration (\( \mu g \text{ C g}^{-1} \text{ day}^{-1} \)) respectively, at 25°C with no nutrients added.

The temperature response (\( Q_{10} \)) value was estimated as:

\[
Q_{10} = \frac{R_2}{R_1}^{\frac{10}{T_2-T_1}}
\]

(2)

Where \( R_1 \) and \( R_2 \) represent the measured respiration rate (\( \mu g \text{ C g}^{-1} \text{ day}^{-1} \)) at temperatures \( T_1 \) and \( T_2 \) (5°C and 25°C).

Laboratory-defined parameters (i.e. growth rate, temperature sensitivity and BGE) were assumed to be the same for all microbial groups. A complete list of parameters and values is presented in Table S3 (Supplementary Information).

### 3. Results

#### 3.1. Laboratory results and model parameters

Bacterial production in untreated soil was estimated at 0.76 \( \mu g \text{ C g}^{-1} \text{ day}^{-1} \) (SD=0.12), and across all nutrient treatments ranged from 0.560 to 2.196 \( \mu g \text{ C g}^{-1} \text{ day}^{-1} \). Nutrient addition led to increased measured production (low = 0.69 \( \mu g \text{ C g}^{-1} \text{ day}^{-1} \) (SD=0.12), medium = 1.09 \( \mu g \text{ C g}^{-1} \text{ day}^{-1} \) (SD=0.53), high = 1.52 \( \mu g \text{ C g}^{-1} \text{ day}^{-1} \) (SD=0.63)), however variability between replicates was also high and production rates from each nutrient treatment were not significantly different from untreated soil (\( P_{\text{low}}=0.99, P_{\text{medium}}=0.70, P_{\text{high}}=0.10 \)). The increased bacterial production was cross-correlated with quadruplicate measurements of biomass from each treatment, and resulting growth rate coefficients (\( I_{\text{maxH}} \)) for all treatments were within a narrow range (0.359 to 0.550 day\(^{-1}\)) and there was no statistically significant difference in growth rates between each nutrient treatment (Fig. 2b) (\( P_{\text{low-medium}}=0.55, P_{\text{medium-high}}=0.49, P_{\text{none-high}}=0.10 \)). The maximum measured growth rate for a single nutrient treatment, thus equating to the parameter \( I_{\text{maxH}} \), was 0.55 day\(^{-1}\). The 95% confidence range for \( I_{\text{maxH}} \) is 0.50 to 0.60 day\(^{-1}\). This value is, to our knowledge, is the first measured rate of bacterial growth from High-Arctic soils, and falls within the lower end of the plausible range established in Bradley et al. (2015) (0.24 – 4.80 day\(^{-1}\)) (Fig. 3a) for soil microbes from a range of laboratory and modelling studies (Frey et al., 2010;
ly-3

chemicals (year 50 and 113) to 18-.

Rhodoferax 3

Phormidium, and was

-437

-436

-435

-434

-433

-432

-431

-430

-429

-428

-427

-426

-425

-424

-423

-422

-421

-420

-419

-418

-417

-416

-415

-414

-413

-412

-411

-410

-409

-408

-407

-406

-405

-404

-403

-402

-401

-400

399

Ingwersen et al., 2008; Knapp et al., 1983; Zelenev et al., 2000; Stapleton et al., 2005; Darrah, 1991;

Blagodatsky et al., 1998; Vandewerf and Verstraete, 1987; Foereid and Yearsley, 2004; Toal et al.,

2000; Scott et al., 1995). For respiration, significantly higher CO2 headspace concentration were

detected in the live incubations at 25°C relative to killed controls (P < 0.05). Average respiration rate at

5°C was 1.61 μg C g⁻¹ day⁻¹ and there was a significant increase in soil respiration at 25°C (12.83 μg C

g⁻¹ day⁻¹) (Fig. 2c) (P < 0.05). The Q₁₀ value for Midtre Lovénbreen forefield soils was thus calculated

as 2.90, and a 95% confidence range was established as 2.65 to 3.16. This was at the upper end of the

plausible range previously identified in Bradley et al. (2015) (Fig. 3b). Based on measured values of

bacterial production and respiration, BGE (Yb) was 0.06, with a 95% confidence range of 0.05 to 0.07

(Fig. 3c). Final calculated values for model parameters are summarized in Table S3 (Supplementary

Information).

The results from microscopy determination of biomass are presented in Table 2. In the freshly exposed

soil (year 0) heterotrophic biomass was low (0.059 μg C g⁻¹), increased substantially to 0.244 μg C g⁻¹

in 29 year old soils, and was an order or magnitude higher (2.00 μg C g⁻¹) in 113 year old soils.

Autotrophic biomass was considerably higher than heterotrophic biomass and increased by roughly an

order of magnitude from year 0 (0.171 μg C g⁻¹) to year 29 (1.07 μg C g⁻¹) and approximately doubled

by year 113 (2.58 μg C g⁻¹). TOC in freshly exposed soil was approximately 0.793 mg C g⁻¹).

16S data was categorized into microbial groups (A1.3 and H1.3) as defined by the model formulation.

Chemolithoautotrophs, such as known iron or sulphur oxidizers (genera Acidithiobacillus, Thiobacillus,

Gallionella, Sulfurimonas) were assigned into the A1 group. Phototrophic microorganisms, such as
cyanobacteria (Phormidium, Leptolyngbya) and phototrophic bacteria (Rhodoferax, Erythrobacter,

Halomicronema) were allocated into group A2, while heterocyst forming cyanobacteria from the orders

Nostocales and Stigonematales were assigned to group the A3 (nitrogen-fixing autotrophs). Members

of the family Comamonadaceae of the Betaproteobacteria are known glacial dwelling microorganisms

(Yde et al., 2010) and were thus included into the group H1. General soil heterotrophic microorganisms

(mainly members of Alphaproteobacteria, Actinobacteria, Bacterioidetes and Acidobacteria) were

assigned into group H2 (general soil heterotrophs). Lastly, group H3 consisted of heterotrophic nitrogen

fixers, mainly Azospirillum, Bradyrhizobium, Devosia, Clostridium, Frankia and Rhizobium. Pathogens,

cyanobacteria from the orders

"Uncategorized" group. Glacial microbes accounted for 43 to 45 % of reads in year 0 and 5, and declined

in older soils (year 50 and 113) to 18 to 22%. The glacial community was predominantly

chemolithoautotrophic (A1). Typical soil bacteria (A2 and H2) increased from low abundance (30% and

40% in years 0 and 5 respectively) to relatively high abundance (63 to 67% of reads) in years 50 and

113. Nitrogen fixing bacteria were prevalent in recently exposed soils (14% in year 0) but low in relative

abundance in soils above 5 years of age (4 to 6% in years 5, 50 and 113). In the freshly exposed soil

(year 0), the microbial community was relatively evenly distributed between heterotrophs (43%) and

autotrophs (44%). In developed soils, the relative abundance of heterotrophs increased (up to 74% of
reads in years 50 and 113). Important to note is the fact that between 8 and 21% of the reads across all samples could not be classified.

### 3.2. Model Results

The model predicted an accumulation of autotrophic and heterotrophic biomass over 120 years (Fig. 4a and 4b). Biomass and nutrient concentrations were initially extremely low (total biomass < 0.25 µg C g⁻¹, DIN < 4.0 µg N g⁻¹, DIP < 3.0 µg P g⁻¹), and biological activity in initial soils was also low (Table 3). There was an order of magnitude increase in total microbial biomass in years 10 to 60. Nitrogen-fixing autotrophs (A₃) and heterotrophs (H₂), and soil heterotrophs (H₂) experienced rapid growth during this period. Glacial and soil autotrophs (A₁-2) and glacial heterotrophs (H₁) remained low. Bacterial production increased by roughly two orders of magnitude (Table 3). Organic carbon (labile and refractory) increased (Fig. 4c), whilst DIN and DIP concentrations increased by approximately an order of magnitude in the first 60 years (Fig. 4d). During the later stages of soil development (years 60 to 120), biomass increased rapidly due to the rapid growth of soil organisms (A₂ and H₂), which outcompeted nitrogen-fixers. The model showed a rapid exhaustion of labile organic carbon (years 50 to 100), while refractory carbon accumulated slowly. Nutrients (DIN and DIP) accumulated at a relatively constant rate. Microbial activity, including bacterial production, nitrogen fixation and DIN assimilation, was high relative to early stages (Table 3).

A carbon budget of fluxes through the substrate pool is presented in Fig. 5. Daily fluxes are presented in panels (a) for year 5, (b) for year 50 and (c) for year 113, and annual fluxes up to year 120 are presented in (d). In recently exposed soils (5 years), allochthonous inputs were the only noticeable carbon flux, outweighing heterotrophic growth and respiration, and the contribution of substrate from necromass and exudates by over two orders of magnitude (Fig. 5a). Thus, the total change in carbon (black line) closely resembled allochthonous input. In the intermediate stages (Fig. 5b), there was substantial depletion from the substrate pool due to heterotrophic activity. Heterotrophic growth (red line) was low despite high substrate consumption and respiration (dark blue line). In the late stages of soil development, the flux of microbial necromass was a significant contributor to the organic substrate pools (Fig. 5c). Carbon fluxes in mid to late stages of soil development were highly seasonal (Fig. 5b and 5c). Biotic fluxes (e.g. respiration) were up to six times higher during the summer (July to September) compared to the winter (November to April), however a base rate of heterotrophic respiration and turnover of microbial biomass was sustained over winter. Figure 4d shows that the contribution of microbial necromass rose steadily throughout the simulation (light blue line), however was not sufficient to compensate the uptake of carbon substrate, thus leading to overall depletion between years 50 to 110 (black line). The contribution of exudates (purple line) to substrate was minimal at all soil ages.

### 4. Discussion

#### 4.1. Determination of parameters and model predictions
Figure 6 illustrates the influence of the site-specific, laboratory-derived parameters on microbial biomass predictions. It compares the range of predicted microbial biomass based on laboratory-determined parameters (yellow) to the entire plausible parameter range (red; Bradley et al. (2015)). Predicted biomass with the average laboratory-derived value is indicated by the black line. For $l_{max}$, predicted biomass with laboratory-derived parameters (yellow shading) was towards the lower end of the plausible range (Fig. 6a) because refined growth rates were significantly lower than the maximum values explored previously. This was mostly due to a significant reduction in autotrophic biomass ($A_{13}$). With high growth rates, there was a sharp early increase in biomass (years 10 to 20) followed by slower growth phase (years 20 to 120). Model results with laboratory-derived growth rates showed that the exponential growth phase occurred later (years 40 to 80) and was more prolonged, but total biomass was considerably lower. There was a substantial reduction in the plausible range in predicted microbial biomass.

There was a substantial reduction in the plausible range in predicted microbial biomass (Fig. 6b) from the measured temperature sensitivity ($Q_{10}$) (yellow) compared to the previous range (red). Soil microbial communities in Polar regions must contend with extremely harsh environmental conditions such as cold temperatures, frequent freeze-thaw cycles, low water availability, low nutrient availability, high exposure to ultraviolet radiation in the summer, and prolonged periods of darkness in winter. These factors profoundly impact their metabolism and survival strategies and ultimately shape the structure of the microbial community (Cary et al., 2010). High $Q_{10}$ values, as derived here, are typical of cold environments and cold adapted organisms and this has been associated with the survival of biomass under prolonged periods of harsh environmental conditions (Schipper et al., 2014). An investigation into the metabolism of microbial communities in biological soils crusts in recently exposed soils from the Austre Broggerbreen Glacier, approximately 6 km from the Midtre Lovénbreen catchment, also derived a high $Q_{10}$ (3.1) (Yoshitake et al., 2010). The Midtre Lovénbreen catchment, in Svalbard, experiences a relatively extreme Arctic climate. The high $Q_{10}$ ultimately lowers the overall rate of biomass accumulation in ultra-oligotrophic soils and a baseline population is maintained.

The low measured BGE (0.06) suggested that a high proportion (94%) of substrate consumed by heterotrophs is remineralized (degrading organic substrate into DIC (CO$_2$), DIN and DIP), with very little incorporated into biomass (6%). Low BGE encouraged the liberation and release of nutrients to the soil and thus the overall growth response of the total microbial biomass was more rapid due to higher soil nutrient concentrations (Fig. 6c). However, due to the low BGE, there was a high rate of substrate degradation, and as such, labile substrate was rapidly depleted when heterotrophic biomass was high (Fig. 4c). Heterotrophic growth requires that a substantial amount of substrate is degraded – thus, although autotrophic production outweighed heterotrophic production at all stages of development (Fig. 4e), the soil was predicted by the model to be a net source of CO$_2$ to the atmosphere over the first 120 years of exposure (Fig. 4f). Heterotrophic growth and respiration (and thus net ecosystem production and carbon fluxes) are strongly dependant on the availability of soil organic carbon. Poorly quantified rates of allochthonous organic carbon deposition and its quality may lead to generally high
uncertainty in the net ecosystem production due to potentially enhanced heterotrophic growth resulting
from higher organic carbon deposition, or lower heterotrophic growth resulting from substrate limitation
in low-deposition scenarios. Soil CO₂ efflux is highly sensitive to variable net ecosystem production,
thus simulated net ecosystem production estimates must be interpreted cautiously until sufficient field
data emerges (e.g. from in situ measurement of soil gas exchange). The calculation of BGE assumes
that bacterial respiration is the major contributor to measured CO₂ gas exchange rates from soil
microcosms. In reality, all active and living soil organisms are likely to contribute to measured CO₂
fluxes, however due to limitations with experimental protocols, it is extremely difficult to determine the
relative contribution of various organisms to total respiration. Microscopy analysis showed limited
presence of fungi and protozoa suggesting that the biological community of the soil community is mainly
bacterial. Nevertheless, by attributing total measured CO₂ fluxes solely to bacteria, BGE may be under-
estimated (due to an overestimation of respiration rates attributed to the bacterial community). Thus,
we cannot exclude that our low BGE values might be in part an artefact of this experimental limitation.
However, although there are very few measurements of BGE in cold glaciated environments, our
estimate of BGE is in good agreement with previous studies, which have suggested values ranging
between 0.0035 and 0.033 (Anesio et al., 2010; Hodson et al., 2007). Therefore, we are confident that
BGE values measured here fall within a realistic range.

Three assumptions are made in the assignment of measured parameters to SHIMMER as applied to
the High-Arctic field site. The first assumption of SHIMMER is that parameter values remain constant
throughout the duration of the simulation. Empirical evidence suggests that parameters defined as fixed
in SHIMMER (e.g. Q₁₀) may be variable over time, however in SHIMMER, like many numerical
modelling formulations, changing environmental (temperature, light) and geochemical (carbon
substrate, available nitrogen, available phosphorus) conditions drive subsequent variability in microbial
activity via mathematical formulations (e.g. Monod kinetics, see Bradley et al. (2015)) affixed to
parameter values. A second assumption is the assignment of measured rates to parameters for all
microbial functional groups. Rather than taxonomic based classification, SHIMMER distinguishes and
classifies microbial communities based on functional traits. These mathematical formulations assigned
to, for example, microbial growth, are different between groups to represent distinct functional traits
associated with that group. Whilst actual rates may be different between different organisms, for the
level of model complexity and outputs required, a community measurement of those parameters is
sufficient, particularly considering that the differences are accounted for in the mathematical formulation
of SHIMMER (see Bradley et al. (2015)). Third, maximum microbial growth rate at T_ref (25°C, Bradley
et al. (2015)) as modelled in SHIMMER is modified by Monod terms that account for nutrient limitation
(e.g. Monod terms), as well as a temperature response function (Q₁₀) to estimate actual growth rate at
ambient temperature. A major objective of of this study was to improve model performance by
constraining previously identified key model parameters (see sensitivity study results in Bradley et al.
(2015)) through specifically designed laboratory experiments. We showed this by comparing model
simulation results applying measured, site-specific parameter with simulation results using a range of
parameter values reported in the literature (Fig. 6).
4.2. Microbial biomass dynamics and community structure

Measured microbial biomass in the initial soils of Midtre Lovénbreen (0.23 µg C g⁻¹, 0 years) was very low compared to initial soils in other deglaciated forefields of equivalent ages in lower latitudes, for example in the Alps (4 µg C g⁻¹) (Bernasconi et al., 2011; Tscherko et al., 2003) and Canada (6 µg C g⁻¹) (Insam and Haselwandter, 1989). However, our microbial biomass values are more similar to other recently deglaciated soils in Antarctica (Ecology Glacier - 0.88 µg C g⁻¹) (Zdanowski et al., 2013). Low biomass is possibly a result of the harsh, ultra-oligotrophic and nutrient limiting environment of the High Arctic and Antarctica, where low temperature and longer winters limit the summer growth phase, especially compared to an Alpine system (Tscherko et al., 2003; Bernasconi et al., 2011).

The initial microbial community structure in our samples was predominantly autotrophic (74.5%). In the years following exposure, we observed an increase in autotrophs and heterotrophs with soil age (Table 2), presumably due to the establishment and growth of stable soil microbial communities (Schulz et al., 2013; Bradley et al., 2014). Both the observations and modelling results suggested that there was no substantial increase in heterotrophic biomass during the initial and early-intermediate stages of soil development (years 0 to 40), which was then followed by a growth phase whereby biomass increased by roughly an order of magnitude. Overall, the model and the microscopy data were in good agreement accounting for the limitations in both techniques, spatial heterogeneity, and the oscillations in biomass arising from seasonality (Fig. 7). SHIMMER predicted that low initial microbial populations have the potential to considerably increase in population density during several decades of soil development. This data thus supports the hypothesis that the observed increase in microbial biomass with soil age is a consequence of in situ growth and activity. The pattern of microbial abundance observed in the Midtre Lovénbreen forefield broadly resembles that of other glacier forefields worldwide (see Bradley et al. (2014)). For example, data from the Rootmoos Ferrer (Austria) (Insam and Haselwandter, 1989), Athabasca (Canada) (Insam and Haselwandter, 1989), Damma (Switzerland) (Bernasconi et al., 2011; Schulz et al., 2013) and Puca (Peru) (Schmidt et al., 2008) glacier forefields find increased microbial biomass and activity over decades to centuries of soil development following exposure.

The genomic data indicated that glacial microbes (such as members of the family Comamonadaceae.) are dominant in recently exposed soils, in agreement with model results (Fig. 8). The community structure in year 5 was heavily dominated by chemolithoautotrophs (A1) (including taxa *Thiobacillus, Rhodoplanes, Acidithiobacillus, Nitrospira, Sulfurimonas* and others), which reflected findings from previous studies whereby chemolithoautotrophic bacteria contribute to the oxidation of FeS₂ in proglacial moraines in Midtre Lovénbreen (Borin et al., 2010; Mapelli et al., 2011). These processes are also commonly described in other subglacial habitats (Boyd et al., 2014; Hamilton et al., 2013). Based on 16S data, the glacial community declined in relative abundance with soil age. This finding was also reflected in the model in years 50 and 113. As the age of the soil progressed, there was typically greater abundance of microbes representing typical soil bacteria (groups A₂ and H₂ including taxa *Geobacter, Micrococcus, Actinoplanes, Sphingomonas, Pedobacter and Devosia, Frankia, Rhizobium*) in the 16S
data and the model, thus the relative abundance of glacial microbes decreased. Relative abundance of microbial communities across the chronosequence is plotted at the phylum and genus level in the Supplementary Information (Fig. S4 and S5). The overall trends show the relative increase in the proportion of Acidobacteria with the soil age. They contain typical soil bacteria and are thus often used as markers of soil formation and soil development. They are usually associated with plant covered older soils with lower pH as they specialize in degradation of plant recalcitrant organic compounds. The younger soils, on the other hand contained relatively higher proportion of sequences of Proteobacteria (particularly Betaproteobacteria), Bacteroidetes and Cyanobacteria, i.e. groups often associated with supra or subglacial habitats.

Microscopic analyses indicated low total biomass in recently exposed soils (up to 1.7 µg C g⁻¹ in soil exposed for 50 years) that was comprised predominantly of autotrophic bacteria. Model simulations agreed well with microscopy derived data. Overall, the 16S data, when categorised into functional groups as defined by the model, agreed well with the microscopy and model output in the very early stages of soil development. However, in later stages of soil development (50 years and older), microscopy and modelling suggested a continuation of predominantly autotrophic soil microbial communities whereas 16S sequence data notably indicated a predominantly heterotrophic community. With extremely low biomass, cell counts derived from microscopy, as well as representation of relative abundance by 16S extraction and amplification, can be largely skewed by relatively small changes in the soil microbial community. Furthermore, the comparative difficulty to lyse autotrophic bacteria (such as some groups of cyanobacteria) from an environmental sample compared to heterotrophic bacteria, and thus successfully amplify the 16S gene during the PCR process, may skew 16S sequence data in favour of heterotrophic sequence reads. Incomplete sample depth (shown by rarefaction curves, Fig S6), the differential extractability of DNA from different organisms, and variation in rDNA copy number adds potential for error to the presented 16S data, and thus absolute numbers should be treated with caution. SHIMMER is an ambitious model in that it attempts to simulate, predict and constrain multiple functional types of bacteria species in a numerical framework. Numerical models containing multiple species or multiple microbial functional groups are often extremely challenging to constrain (Servedio et al., 2014; Hellweg and Bucci, 2009; Jessup et al., 2004; Larsen et al., 2012), and as such, the majority of microbial soil models often only resolve one or two living biomass pool that represents the bulk activity and function of the entire community (see e.g. Manzoni et al. (2004), Manzoni and Porporato (2007), Blagodatsky and Richter (1998), Ingwersen et al. (2008), Wang et al. (2014) and others). Our rationale for resolving six distinct functional groups was to quantitatively assess, using modelling, the relative importance and role of each functional group at different stages of soil development. Regardless of discrepancies in older soils (over 50 years since exposure), both the 16S and microscopy data indicated that there was a mixed community of autotrophs and heterotrophs in soils of all ages, which was supported by modelling, since no functional groups were extirpated over simulations representing 120 years of soil development. Thus, SHIMMER is able to capture the diversity of the samples over 120 years of soil development, but the detailed community composition requires further investigation.
Nitrogen-fixing bacteria such as *Nostoc*, *Rivularia*, *Pseudanabaena* and *Rhodobacter* were prevalent in recently exposed soils but declined in relative abundance with soil age. By fixing N\textsubscript{2} instead of assimilating DIN, the model predicted that nitrogen-fixers were able to grow rapidly in the early stages relative to other organisms (Fig. 4a, 4b). The model prediction supports findings by previous studies demonstrating the importance of nitrogen fixation in Alpine (Duc et al., 2009; Schmidt et al., 2008) and Antarctic (Strauss et al., 2012) glacier forefields and other High-Arctic (Svalbard, Greenland) glacial ecosystems (Telling et al., 2011; Telling et al., 2012). However, there was poor agreement on the relative abundance of nitrogen-fixers between the model and the 16S data in the later stages of soil development (years 50 to 120), particularly between autotrophs and heterotrophs. The model over-predicted the relative abundance of nitrogen-fixing organisms (Fig. 8). The majority of the biomass of the autotrophic nitrogen-fixers was composed of sequences belonging to the cyanobacterium from the genus *Nostoc*. *Nostoc* forms macroscopically visible colonies that grow on the surface of the soils. Its distribution in the Arctic soils is thus extremely patchy and therefore, part of the discrepancy between the 16S data and the model regarding the relative distribution of the A\textsubscript{3} group in the older soils could be due to under-sampling of the *Nostoc* colonies as a consequence of a random sampling approach. Furthermore, allochthonous inputs of nitrogen to the Arctic (e.g. aerial deposition (Geng et al., 2014)) strongly affect the productivity of microbial ecosystems and the requirement of nitrogen fixation for microbes (Bjorkman et al., 2013; Kuhnel et al., 2013; Kuhnel et al., 2011; Hodson et al., 2010; Telling et al., 2012; Galloway et al., 2008). Thus, uncertainty in the allochthonous availability of nitrogen strongly affects nitrogen fixation rates. In attempting to replicate a qualitative understanding of the nitrogen cycle in a quantitative mathematical modelling framework, the predicted importance of nitrogen-fixing organisms may be over-estimated. The poor agreement in the relative abundance of nitrogen-fixers between the model and the 16S data indicates an incomplete understanding of allochthonous versus autochthonous nutrient availability. Allochthonous nutrient availability is a known source of uncertainty (Bradley et al., 2014; Schulz et al., 2013; Schmidt et al., 2008), and addressing this concern is the subject of future work.

16S data is an exciting resource of information that is rarely (or never) used to test numerical process-based biogeochemical models. However, the environment (difficulty to extract DNA), the presentation (percentages of low concentration and thus easy to shift relative abundance), the potentially high proportion of dead or dormant cells (which may be present in sequence data but are not necessarily metabolically active), the differential extractability of DNA from different organisms and variation in rDNA copy number, and uncertainties in model formulation make comparisons challenging. In making this first attempt at comparison of model output to 16S data, we hope to spark discussion and further development of approaches that have similar objectives in order to improve future model performance.

**4.3. Net ecosystem metabolism and carbon budget**
Allochthonous carbon inputs were the most significant contributor to recently exposed soils (e.g. year 5), since the total change in substrate closely followed this flux (Fig. 5). In older soils (year 113), biotic fluxes were substantially higher, and microbial necromass contributed equally as a source of organic substrate compared to allochthonous deposition. In the older soils, heterotrophic growth and respiration caused substantial consumption and thus depletion of available carbon stocks. This evidence thus supports the hypothesis that carbon fluxes in very recently exposed soils are low and are dominated by abiotic processes (i.e. allochthonous deposition), whereas biotic processes (such as microbial growth, respiration and cell death) play a greater role in developed soils with increased microbial abundance and activity. These findings for the Midtre Lovénbreen glacier in the High-Arctic, are similar to what has been observed based on empirical evidence from Alpine settings (at the Damma Glacier, Switzerland (Smittenberg et al., 2012; Guelland et al., 2013)).

The seasonality of carbon fluxes predicted by the model (Fig. 5b and 5c) related to the high measured $Q_{10}$ values. High seasonal variation in biotic fluxes and rates is typical of cryospheric soil ecosystems (Schostag et al., 2015) including Alpine glacier forefield soils (Lazzaro et al., 2012; Lazzaro et al., 2015). However, microbial activity has been shown to persist during winter under insulating layers of snow and in sub-zero temperatures (Zhang et al., 2014). Modelling also predicted sustained organic substrate degradation, microbial turnover and net heterotrophy during the winter (Fig. 5b and 5c), as documented in other glacier forefield studies from an Alpine setting (Guelland et al., 2013b).

The low measured BGE has three important consequences. Firstly, low BGE suggests that a large pool of substrate is required to support heterotrophic growth. Low-efficiency heterotrophic growth lead to the rapid depletion of substrate; therefore high allochthonous inputs were required to maintain a sizeable pool. In older soils (years 80 to 120), increased inputs from microbial necromass (blue line, Fig. 5d) sustained substrate supply to heterotrophs. The sources of allochthonous carbon substrate to the glacier forefield include meltwater inputs derived from the supraglacial and subglacial ecosystems (Stibal et al., 2008; Hodson et al., 2005; Mindl et al., 2007), snow algae (which are known to be prolific primary colonizers and producers in High Arctic snow packs (Lutz et al., 2015; Lutz et al., 2014), atmospheric deposition (Kuhnel et al., 2013) and ornithogenic deposition (e.g. faecal matter of birds and animals) (Jakubas et al., 2008; Ziolek and Melke, 2014; Luoto et al., 2015; Michelutti et al., 2009; Michelutti et al., 2011; Moe et al., 2009). Microbial dynamics are moderately sensitive to external allochthonous inputs of substrate (Bradley et al., 2015), and addressing the uncertainty associated with this flux is an important question to address in future research.

Secondly, low BGE causes a net efflux of CO$_2$ over the first 120 years of soil development despite high autotrophic production (Fig. 4e and 4f). Recent literature has explored the carbon dynamics of glacier forefield ecosystems, finding highly variable soil respiration rates (Bekku et al., 2004; Schulz et al., 2013; Guelland et al., 2013a). Future studies should focus on quantifying carbon and nutrient transformations and the potential for forefield systems to impact global biogeochemical cycles in response to future climate change (Smittenberg et al., 2012) and in the context of large-scale ice retreat.
Thirdly, high rates of substrate degradation encouraged by low BGE were responsible for rapid nutrient release. Modelling suggested that microbial growth was strongly inhibited by low nutrient availability in initial soils (4 µg N g\(^{-1}\), 2 to 10 µg P g\(^{-1}\)) (Fig. 4d). This is consistent with findings from the Hailuogou Glacier (Gongga Shan, China) and Damma Glacier (Switzerland) (Prietzel et al., 2013). Low BGE is predicted by the model to have a very important role in encouraging the release of nutrients from organic material more rapidly, thereby increasing total bacterial production in the intermediate stages of soil development. Increased nutrient availability with increased heterotrophic biomass is consistent with recent observations from glacier forefields (Bekku et al., 2004; Schulz et al., 2013; Schmidt et al., 2008).

5. Conclusions

We used laboratory-based mesocosm experiments to measure three key model parameters: maximum microbial growth rate \(I_{\text{max}}\) (by incorporation of \(^3\)H-leucine), BGE \(Y\) (by measuring respiration rates) and the temperature response \(Q_{10}\) (by measuring rates at different ambient temperatures). Laboratory-derived parameters were comparable with previous estimations. We refined model predictions constraining site-specific parameters by lab experiments, thus decreasing parameter uncertainty and narrowing the range of model output over nominal environmental conditions. A comparison of model simulations using laboratory-derived parameter values and previously defined parameter values showed that the coupling of high \(Q_{10}\) values and low BGE were important factors in controlling biomass accumulation due to promoting survival of biomass during periods of low temperature, and the enhanced recycling of nutrients through organic matter degradation, respectively.

Our results demonstrated that \textit{in situ} microbial growth lead to the overall accumulation of microbial biomass in the Midtre Lovénbreen forefield during the first century of soil development following exposure. Furthermore, carbon fluxes increased in older soils due to elevated biotic (microbial) activity. Microbial dynamics at the initial stages of soil development in glacial forefields do not contribute to significant accumulation of organic carbon due to the very low growth efficiency of the microbial community, resulting in a net efflux of CO\(_2\) from those habitats. However, the low bacterial growth efficiency in glacial forefields is also responsible for high rates of nutrient remineralization that most probably has an important role on the establishment of plants at older ages. The relative importance of allochthonous versus autochthonous substrate and nutrients is the focus of future research.

This exercise shows how an integrated model-data approach can improve understanding and predictions of microbial dynamics in forefield soils and disentangle complex process interactions to ascertain the relative importance of each process independently. This would, for annual budgets, be extremely challenging with a purely empirical approach. Nevertheless, more clarity and data are needed in tracing the dynamics and interactions of these carbon pools to improve confidence and validate model simulations. Proglacial zones are expanding due to accelerated ice retreat. Thus, glacier forefields are becoming an increasingly important novel habitat for microorganisms in glaciated regions experiencing rapid changes in climate. This combined approach explored detailed microbial and biogeochemical
dynamics of soil development with the view to obtaining a more holistic picture of soil development in a warmer and increasingly ice-free future world.

Acknowledgements

We thank Siegrid Debatin, Marion Maturilli, and Julia Boike (AWI) for support in acquiring meteorological and radiation data, Simon Cobb, James Williams, Jane Coghill and Christy Waterfall (University of Bristol) for laboratory assistance, and Nicholas Cox and James Wake for assistance in the field and use of the UK Station Arctic Research base in Ny-Ålesund. We also thank Arwyn Edwards (Aberystwyth University) and three anonymous referees who provided valuable comments on the manuscript. This research was supported by NERC grant no. NE/J02399X/1 to A. M. Anesio. S. Arndt acknowledges support from NERC grant no. NE/I021322/1.

References


Hodson, A., Anesio, A. M., Ng, F., Watson, R., Quirk, J., Irvine-Fynn, T., Dye, A., Clark, C., McCloy, P., Kohler, J., and Sattler, B.: A glacier respires: Quantifying the distribution and
respiration CO2 flux of cryoconite across an entire Arctic supraglacial ecosystem, J Geophys Res-Biogeo, 112, Artn G04s36
Doi 10.3402/Polar.V32i0.19136, 2013.


Figure 1. Midtre Lovénbreen glacier and forefield in Svalbard, the location of sampling sites and approximate age of soil.
Figure 2. Measurements of (a) bacterial carbon production and (b) growth rate, derived from $^3$H-leucine assays at different nutrient conditions, and (c) bacterial respiration at 5°C and 25°C.
Figure 3. A comparison of previously established ranges for parameters (Bradley et al., 2015) with laboratory-derived values for (a) maximum growth rate ($I_{\text{max}}$), (b) temperature response ($Q_{10}$), (c) BGE ($Y$).
Figure 4. Modelled (a) autotrophic biomass, (b) heterotrophic biomass, (c) carbon substrate, (d) nutrients, (e) bacterial production and (f) net ecosystem production, with laboratory-derived parameter values.
Figure 5. Illustration of daily carbon fluxes for (a) 5, (b) 50 and (c) 113 year old soil, and (d) annual carbon flux over 120 years. Microbial necromass (light blue), exudates (purple) and allochthonous sources (yellow) contribute to the substrate pool (black), and heterotrophic growth (red) and respiration (dark blue) deplete it.
Figure 6. A comparison of predicted microbial biomass with laboratory-derived parameter values (yellow) and previously established parameter values (Bradley et al., 2015) (red) for variation in the following parameters: (a) maximum growth rate ($l_{max}$), (b) temperature response ($Q_{10}$), (c) BGE ($\gamma$).
Figure 7. Model predictions of (a) autotrophic and (b) heterotrophic biomass (black line), compared to observational data (red) derived from microscopy.
Figure 8. A comparison of microbial diversity from model output and genomic analyses at 0 year old, 5 year old, 50 year old and 113 year old soil.
**Table 1. State variables and initial values.**

<table>
<thead>
<tr>
<th>State Variable</th>
<th>Units</th>
<th>Description</th>
<th>Initial value (year 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Glacial chemolithoautotrophs</td>
<td>0.0547</td>
</tr>
<tr>
<td>$A_2$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Soil autotrophs</td>
<td>0.0266</td>
</tr>
<tr>
<td>$A_3$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Nitrogen fixing soil autotrophs</td>
<td>0.0355</td>
</tr>
<tr>
<td>$H_1$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Glacial heterotrophs</td>
<td>0.0576</td>
</tr>
<tr>
<td>$H_2$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Soil heterotrophs</td>
<td>0.0530</td>
</tr>
<tr>
<td>$H_3$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Nitrogen fixing soil heterotrophs</td>
<td>0.0025</td>
</tr>
<tr>
<td>$S_1$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Labile organic carbon</td>
<td>291.895</td>
</tr>
<tr>
<td>$S_2$</td>
<td>$\mu g$ C $g^{-1}$</td>
<td>Refractory organic carbon</td>
<td>681.089</td>
</tr>
<tr>
<td>DIN</td>
<td>$\mu g$ N $g^{-1}$</td>
<td>Dissolved inorganic nitrogen (DIN)</td>
<td>3.530</td>
</tr>
<tr>
<td>DIP</td>
<td>$\mu g$ P $g^{-1}$</td>
<td>Dissolved inorganic phosphorus (DIP)</td>
<td>2.078</td>
</tr>
<tr>
<td>ON$_1$</td>
<td>$\mu g$ N $g^{-1}$</td>
<td>Labile organic nitrogen</td>
<td>41.157</td>
</tr>
<tr>
<td>ON$_2$</td>
<td>$\mu g$ N $g^{-1}$</td>
<td>Refractory organic nitrogen</td>
<td>96.034</td>
</tr>
<tr>
<td>OP$_1$</td>
<td>$\mu g$ P $g^{-1}$</td>
<td>Labile organic phosphorus</td>
<td>24.227</td>
</tr>
<tr>
<td>OP$_2$</td>
<td>$\mu g$ P $g^{-1}$</td>
<td>Refractory organic phosphorus</td>
<td>56.530</td>
</tr>
</tbody>
</table>
Table 2. Microbial biomass in the forefield of Midtre Lovénbreen (brackets show 1 standard deviation)

<table>
<thead>
<tr>
<th>Soil Age (years)</th>
<th>Autotrophic biomass ($\mu g \text{ C g}^{-1}$)</th>
<th>Heterotrophic biomass ($\mu g \text{ C g}^{-1}$)</th>
<th>Total Organic Carbon ($\mu g \text{ C g}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.171 (0.042)</td>
<td>0.039 (0.034)</td>
<td>792.984 (127.206)</td>
</tr>
<tr>
<td>3</td>
<td>0.287 (0.155)</td>
<td>0.064 (0.029)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.561 (0.143)</td>
<td>0.083 (0.065)</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>1.072 (0.487)</td>
<td>0.244 (0.142)</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>1.497 (0.601)</td>
<td>0.197 (0.184)</td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>2.581 (0.927)</td>
<td>2.000 (0.885)</td>
<td></td>
</tr>
<tr>
<td>Soil Age (years)</td>
<td>Autotrophic biomass (µg C g(^{-1}))</td>
<td>Heterotrophic biomass (µg C g(^{-1}))</td>
<td>Autotrophic production (µg C g(^{-1}) y(^{-1}))</td>
</tr>
</tbody>
</table>
|-----------------|----------------------------------------|-----------------------------------------|-------------------------------------|-------------------------------------|---------------------------------------|---------------------------------|---------------------------------
| 0               | 0.117                                  | 0.111                                   | 0.002                               | 0.001                               | -0.011                                | 2.0 \times 10^{-4}               | 2.0 \times 10^{-4}               |
| 3               | 0.117                                  | 0.105                                   | 0.003                               | 0.001                               | -0.020                                | 3.0 \times 10^{-4}               | 3.0 \times 10^{-4}               |
| 5               | 0.119                                  | 0.102                                   | 0.004                               | 0.001                               | -0.025                                | 4.0 \times 10^{-4}               | 4.0 \times 10^{-4}               |
| 29              | 0.359                                  | 0.147                                   | 0.050                               | 0.012                               | -0.391                                | 0.002                           | 0.006                            |
| 50              | 0.860                                  | 0.591                                   | 0.187                               | 0.113                               | -4.311                                | 0.022                           | 0.021                            |
| 113             | 4.414                                  | 1.331                                   | 3.093                               | 0.376                               | -4.031                                | 0.458                           | 0.031                            |