Lipid load triggers migration to diapause in Arctic Calanus copepods - insights from underwater imaging

Moritz Schmid and Louis Fortier
Takuvik Joint International Laboratory, Québec-Océan, Département de biologie, Université Laval, Québec, Canada
Arctic marine food web

* = kJ x m⁻² x yr⁻¹, Welsh et al. 1992

Darnis et al. 2012
• Important adaptions of zooplankton
  – Diel vertical migration
  – Diapause based on accumulated lipids

  – **Coarse vertical** resolution of *copepod distributions* and lipids
  – Fine scale information relevant for individual-based processes
From nets to imaging

LOKI

1) Plankton concentration net

2) LOKI computer with SSD drive and sensors:
   - Temperature
   - Conductivity
   - Oxygen
   - Fluorescence
   - Pressure

3) LOKI camera

4) Battery

Cod end attached to outflow of camera
- **Calanus hyperboreus**
  - True herbivore, major lipids, diapause

- **C. glacialis**
  - Herbivore with omnivore proclivity, lipids, diapause

- **Metridia longa**
  - Omnivore, little lipids, feeds and reproduces year round
Automatic identification of LOKI images using AI

- 114 categories (Schmid et al. 2016, Methods in Oceanography)
Lipid measuring

2 measures:
a) Mg of total lipids based on area of lipid sac (power law equation), Vogedes et al. (2010)
b) Fullness index (area of lipid sac divided by area of prosome)
Study area

- CCGS *Amundsen*, ArcticNet cruise 2013

- 24h Lagrangian drift, 15/16 August, following surface drifter
Results

C. hyperboreus

Fluorescence (mg m$^{-3}$)

Depth (m)

Abundance (m$^{-3}$)
C. glacialis
C5 population less deep and vertically migrating

M. longa

M. longa and C. glacialis C1-C3 stages rested in surface waters while C. hyperboreus C3 was vertically migrating
Lipids

![Box plot showing lipid fullness for different taxa](image)

- C. hyperboreus
- C. glacialis
- M. longa

Lipid fullness (%)

Taxa:
- C3
- C4
- C5
- F

N=822
Lipid fullness, vertically stratified

C. hyperboreus

Significant difference between day and night

Large parts of the population at depth
Significant difference between day and night

C. glacialis

All individuals at depth
Discussion

– One of the most detailed automatic zooplankton identification models
– Excellent vertical detail of abundances
  → Partially consistent with other studies

• Total lipids and lipid fullness:
  – 80% lipid fullness in *Calanus* females, higher than previously reported
  – Likely first lipid data for *C. hyperboreus* C3 and *M. longa* C5
  – Lipid fullness reflects different life-cycle strategies. 36% fullness in *Metridia*, suggest it can thrive on mixed strategy, buffer substantial food shortage (*sensu* Ashjian et al., 1995)
– All three species showed individuals with higher lipid fullness were deeper. Differing reasons for Metridia (predators) and Calanus (predators and diapause)
– C. glacialis C4/C5 and C. hyperboreus C4; significant differences in lipid distributions between day and night → individuals with lower lipid fullness migrated back to the surface for feeding while individuals with high lipids stayed at depth → diapause
– Transition from DVM to diapause
– C. hyperboreus C4 (~50% lipid fullness) and C. glacialis C5 (~60% lipid fullness), ready for, or in diapause
  • lipid-accumulation-window (LAW) hypothesis (e.g. Maps et al., 2010)
  • Pond and Tarling (2011) Calanoides acutus
Bioenergetic model (Maps et al. 2014) showed *Calanus* females could diapause for over 365 days ➔ spawning

- Also confirms that *C. hyperboreus* C3 and *C. glacialis* C4 did not have enough lipids for diapause

- LOKI images make for an effective way of studying lipids
- Likely most detailed insight into vertical distribution of lipids
We want to thank the crew and officers of the CCGS *Amundsen* without whom this project would not be possible.

Thanks to Frédéric Maps for advice on copepod lipids and applying lipid data from this study to his bioenergetic model. Further thanks to everybody who helped sampling: Cyril Aubry, Maxime Geoffroy, Marianne Falardeau, Mathieu LeBlanc, Catherine Lalande and Alexis Burt. Many thanks to Cyril Aubry for valuable input throughout the project.

We are grateful to all our funding sources, primarily the Canada Foundation for Innovation (CFI) and ArcticNet, a Network of Centres of Excellence of Canada.

Auel H, Klages M, Werner I (2003) Respiration and lipid content of the Arctic copepod Calanus hyperboreus overwintering 1 m above the seafloor at 2,300 m water depth in the Fram Strait. Mar Biol 143:275–282


The image contains several graphs showing fluorescence and abundance data for different depths and times. The graphs are labeled with different times and photon fluxes, such as 6:40h, 102µE, 10:30h, 750µE, 18:45h, 380µE, 23:20h, 45µE, and 2:40h, 18µE. The depth ranges from 0 to 1000 meters, and the fluorescence and abundance are measured in mg m⁻³ and abundance in m⁻³, respectively.

- **C. glacialis**
- **Fluorescence (mg m⁻³)**
- **Depth (m)**
- **Abundance (m⁻³)**

The graphs show variations in fluorescence and abundance at different depths and times, indicating changes in environmental conditions or biological activity.
Note on climate change:
Higher temp $\rightarrow$ increases metabolic demand $\rightarrow$ shorter diapause length (Maps et al. 2014).
However, temp of deeper Arctic waters (> 200 m) not expected to increase to levels affecting diapause in near future (Seidov et al., 2014).
But mismatch scenarios earlier in the season (Falk-Petersen et al., 2007), could disrupt lipid accumulation and lead to starvation (Auel et al., 2003). M. Longa would likely fare better with climate change
Methods

Study area: Greater North Water (NOW) Polynya region
Taxonomic classification system:

114 categories for all plankton and particles
- including subclasses and orientations

Different *C. glacialis* females:
- a) lateral
- b) dorsal short
- c + d) dorsal long
- e + f) antenna in front
Developing the automatic identification model

- 14558 images manually identified (7252 for training, 7306 for model internal testing)
- 52 parameters extracted from images (e.g. length, circularity and greyness level)
Ensemble model of randomly generated decision trees (Cutler et al. 2007, Hochachka et al. 2007)

Machine learning, sub-discipline of artificial intelligence

Start

Length > 6 mm

Area > 10 mm²

C. hyperboreus female

binary split at nodes
### Random set of Columns
(Explanatory variables/Predictors)

<table>
<thead>
<tr>
<th></th>
<th>length</th>
<th>width</th>
<th>area</th>
<th>circularity</th>
<th>greyness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random tree 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random tree 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Random set of Rows
(Observations/Cases)

... and 47 more

Random sampling with replacement
How does the model identify the species on an image?

Random Forests ensemble model
- here only 3 of 10,000 trees

C. glacialis
C. hyperboreus
C. glacialis

Final classification voted
C. glacialis

1 used with permission of http://blog.yhathq.com
Model validations

Internal
  – Based on withheld dataset of 7306 images

External
  – Based on completely and automatically identified stations 101 and 126
  a) 50% of images verified for validation measures
  b) comparison of biological sample (IDed under microscope) vs automatic identifications, n=1000 per station

Original 114 categories (with all orientations) recombined to 63 categories
- C. glacialis C1 dorsal & C. glacialis C1 lateral → C. glacialis C1
Results

**Probability of detection (PD)** (% correct identifications/total manual identifications)

**Specificity (SP)** (% correct identifications/total classifier identifications)

<table>
<thead>
<tr>
<th>Confusion Matrix</th>
<th>Taxa 1</th>
<th>Taxa 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxa 1</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td>Taxa 2</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>Total manual</td>
<td>55</td>
<td>35</td>
</tr>
<tr>
<td>identifications</td>
<td>81%</td>
<td>86%</td>
</tr>
</tbody>
</table>

PD = How good is the machine at detecting the true positives that I manually classified?

SP = How good is the machine at ruling out false positives?

**Model internal validation**

→ PD & SP 86%
→ mean probability of misclassification 2.9%
Confusion matrix, model internal
External validation

a) Based on completely identified stations

Station 101)
- 4820 images, PD and SP 85% each
- mean probability of misclassification 3.1%

Station 126)
- 2163 images, PD and SP 81% each
- mean probability of misclassification 5.2%

**Probability of detection (PD)** (% of correct identifications/total manual identifications)

**Specificity (SP)** (% of correct identifications/total classifier identifications)
Confusion matrix, external station 101
Confusion matrix, external station 126
b) Biological sample vs automatic identification
From zooplankton nets to imaging

Since 1950s, development of imaging systems (Foote 2000, Jaffe et al. 2001)

Now many different exist, e.g. VPR (Woods Hole), UVP (CNRS, Villefranche-sur-Mer)

VPR

UVP

C.S. Davis

D. Luquet
C. hyperboreus

Females

Fullness ratio

> 90% !
Our system:
The **Lightframe Onsight Keyspecies Investigation (LOKI)** system developed by AWI in Germany, newest generation (Schulz et al. 2010)

<table>
<thead>
<tr>
<th>VPR images</th>
<th>UVP 5 images</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

Moeller et al. 2012

Picheral et al. 2010
How does the frequency distribution of lipids look like?

**C. hyperboreus**

- Fullness ratio: rather unimodal

- Total lipids (mg)
  - C3
  - F

**C. glacialis**

- Fullness ratio: rather bimodal

**M. longa**

- Fullness ratio: > 90%
Variables that LOKI browser is extracting from the objects:

- The area of the object in pixel
- The area of the object in mm²
- The shape factor circularity with values of $\geq 1.0$. A value of 1 resembles a circle (Wojnar et al. 2000)
- Convexity (Berger 1990)
- Homogeneity, histogram of the cooccurrence of grey values in an image (Haralick et al. 1973, Knauer and Meffert 2010)
- The mean gray value of pixels within the object frame
- Kurtosis, ‘peakedness’ of the histogram of grey values (Dodge 2003)
- Skewness, grey value distribution (von Hippel 2005)
- Hu moments 1-7, so called moments (moment = a certain particular weighted average of the image pixels' intensities) (Hu 1962)
- Fourier descriptors 1-10, contour based shape descriptors based on fourier transformations (Zhang and Lu 2002)
Obtaining the zooplankton thumbprint (①)

1)

2)

3) 52 parameters extracted

4)
Flowchart showing image analysis and automatic taxonomic identification

- Raw LOKI images
- Taxonomic identification for 5-10% of images
- Completely identified dataset
- Training model
- Rest of unidentified images
- Machine learning software
- Training dataset
Bioenergetic model: Maps et al. (2014)

\[ D = \frac{4}{-M} \left( C_{out}^{1/4} - C_{in}^{1/4} \right) \exp \left( -E \frac{T - T_0}{kTT_0} \right) \]

\[ C_{out} = \frac{2}{3} (1 - LF) \frac{\lambda TL}{LF} \]

\[ C_{in} = \lambda TL + C_{out} \]

\( M (\mu g \ C^{1/4} \ s^{-1}) = \) metabolic constant,
\( E (eV) = \) the activation energy for dormant individuals,
\( T (K) = \) the in situ temperature,
\( T_0 = \) reference temperature (273 K),
\( k = \) Boltzmann constant (8.62 \( 10^{-5} \) eV),
\( \lambda = \) carbon to lipid conversion factor (0.81, based on Kattner and Graeve (1991))