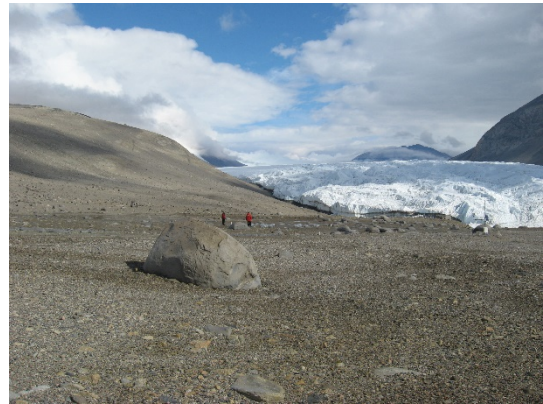


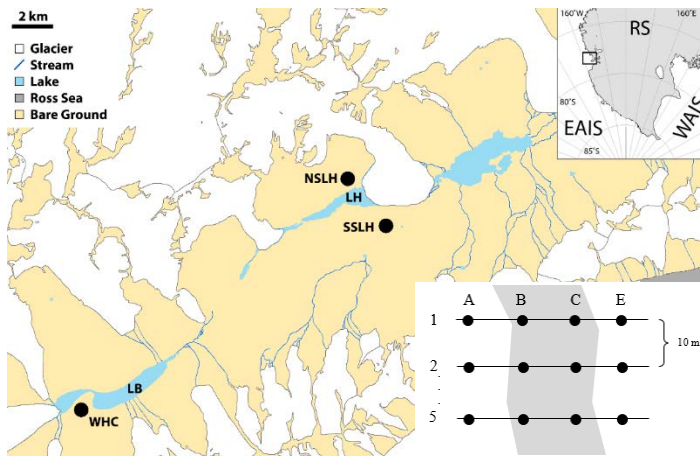
Dataset context:

The McMurdo Dry Valleys in East Antarctica are one of the most extreme desert environments on Earth. Low temperatures and the scarcity of liquid water limit biological processes, and the organisms living here experience conditions near the ‘limits of life’. Vascular vegetation is absent, primary production is very low, and the relatively simple soil biological community is entirely microscopic. Because of the harsh environment, biotic abundance is greatest near sources of liquid water, such as the ephemeral streams that move glacial meltwater from high elevation to the ice-covered lakes during the brief summer season.



However, another underground “cryptic” transfer of water also exists, in the form of water tracks. Water tracks are channels of shallow groundwater that route water downslope through the thawed soil above the permafrost, without surface flow. The additional water available inside water tracks will potentially stimulate soil biology, but they also transfer high volumes of salt (typically 5-10X saltier than off-track soils) that could create osmotic stress for the organisms. Such alterations to soil moisture and geochemistry constitute a change in habitat suitability for soil biology and therefore their ability to mineralize carbon through respiration. Will the enhanced moisture stimulate biological

activity (to increase C mineralization), or will the associated salinity inhibit it? To answer this question, CO₂ flux was measured from water track soils. Notably, soil biota are not the only source of soil CO₂ flux, and the liquid water might also stimulate geochemical fluxes of CO₂, allowing water tracks to also influence the relative importance of biotic vs. geochemical CO₂ fluxes. We hypothesized that inside the water tracks, soils would be wetter but also more saline, and that on balance, this would have a net negative impact on soil CO₂ flux from water track soils.



Three water tracks were sampled: one on the south side of Lake Hoare (SSLH), one on the north side (NSLH), and one in the Lake Bonney basin, affectionately named Wormherder Creek (WCH). Each was sampled in a grid, where 5 transects were sampled at four positions: two inside the water track (visibly wetted surface zone) and two outside the water track (one on each side in dry, light-toned soil). At SSLH, there were five positions, because of its larger width. CO₂ flux from soil was measured at each point. In addition,

chemical parameters often associated with respiration were measured, including soil pH, electrical conductivity (a correlate for salinity), soil water content, and carbon and nitrogen content.

For more information, this research is published in: Ball, Becky A. & Joseph Levy. 2015. The role of water tracks in altering biotic and abiotic soil properties and processes in a polar desert in Antarctica. *Journal of Geophysical Research: Biogeosciences* 120.270-79.

Data description:

The data collected in this project includes:

SampleID (A unique code for each sample)

SiteID (NSLH, SSLH, or WHC)

SiteName (Northside Lake Hoare, Southside Lake Hoare, Wormherder Creek)

Transect (1-5, i.e. replicate samples along the water track)

PosID (Position ID: A-E, positions outside and inside across the water track)

PosDesc (Whether the position is inside or outside the water track, either in-track or off-track)

Flux (CO₂ flux measured from soil, either by respiration or geochemical degassing, in µg C/m²/s)

Temp (Soil temperature, in °C)

SWC (Soil Water Content, in %g water per g soil)

pH (alkalinity of the soil, unitless)

EC (Electrical Conductivity, in µS/cm)

Nitr (Nitrogen content of the soil, in %)

Carb (Total carbon content of the soil, in %)

Corg (Organic carbon content of the soil, in %)

In an Excel spreadsheet, the data look like this, with a total of 65 samples:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	SampleID	SiteID	SiteName	Transect	PosID	PosDesc	Flux	Temp	pH	EC	SWC	Nitr	Carb	Corg
2	SSLH1A	SSLH	Southside	1	A	off-track	0.04	12.45	9.45	438.3	0.50	0.0027	0.44	0.04
3	SSLH1B	SSLH	Southside	1	B	in-track	0.02	12.14	6.5	1655.4	4.82	0.0014	0.13	0.04
4	SSLH1C	SSLH	Southside	1	C	in-track	0.04	12.03	6.79	1702.8	5.10	0.0016	0.08	0.03
5	SSLH1D	SSLH	Southside	1	D	in-track	0.03	9.08	7.22	488.9	12.00	0.0023	0.12	0.04
6	SSLH1E	SSLH	Southside	1	E	off-track	0.01	7.37	7.38	139.1	13.45	0.0033	0.04	0.04
7	SSLH2A	SSLH	Southside	2	A	off-track	0.12	12.7	9.8	91.1	1.94	0.0051	0.15	0.05
8	SSLH2B	SSLH	Southside	2	B	in-track	0.03	10.78	7.59	4919.7	1.94	0.0019	0.19	0.03
9	SSLH2C	SSLH	Southside	2	C	in-track	0.03	9.95	7.88	4094.6	4.00	0.0016	0.18	0.03
10	SSLH2D	SSLH	Southside	2	D	in-track	0.07	10.25	6.96	947.8	4.06	0.0012	0.10	0.03
11	SSLH2E	SSLH	Southside	2	E	off-track	0.02	10.52	8.22	2419.6	1.42	0.0013	0.12	0.02
12	SSLH3A	SSLH	Southside	3	A	off-track	0.09	10.41	9.53	1164.4	2.35	0.0051	0.40	0.06

Database design

1. Identify any redundancy in the spreadsheet.
2. Identify anomalies that may occur due to the redundancy.
3. If there is redundancy, propose a breakdown into multiple spreadsheets/tables to remove it. Draw the spreadsheets/tables with the data.

4. For each spreadsheet/table in the breakdown, identify its primary key. (A primary key is a column or a set of columns that will always identify a unique row in the spreadsheet/table even if the data in the cells change.)

5. For each spreadsheet/table in the breakdown, identify its foreign key(s), if any. (A foreign key is a column or combination of columns referring to values that uniquely identify a row in another spreadsheet/table.)

6. For each table, create a question/query that can be answered using only that table.

7. For each primary-foreign key relationship, create a question/query that can be answered using the relationship between the tables.

8. Draw the design, which is called a schema, of this database, indicating primary and foreign keys.

Database implementation

Now that you have designed and diagrammed the database, we can use database software, such as Microsoft Access, to build the database and conduct queries on the data. To continue with the assignment, open the database file WaterTrackExample.accdb to conduct the queries.

NOTE: For many scientists, our data are initially stored as a large spreadsheet in a program such as Excel. The first step to databasing is to import that spreadsheet into database software and set up the database according to your planned design. To see how this is done using the example data in this exercise, see the tutorial at: <https://youtu.be/vr10il5cfgE>.

Write SQL statements to answer each of the following queries. Note that the first step in answering a query is analyzing the design to see what tables are needed to answer the query. After writing the SQL statements, conduct the query in the Access database to verify that your SQL statements give the correct answer.

NOTE: A very brief tutorial with the basics of how to run a query in Access can be found at: <https://youtu.be/6TXA5TE3u-c> A more detailed primer of querying in both Access and SQL can be found at: <http://databasesmanymajors.faculty.asu.edu/QueryDB/QueryDBAccessSQL.pdf>

- a) Which samples had SWC less than 10%?
- b) Is position D in- or off-track?
- c) What is the full name of the water track abbreviated WHC?
- d) What is the average CO₂ flux of in-track soils?
- e) What is the average temperature at each site?
- f) What is the average pH of the off-track soils at Wormherder Creek?