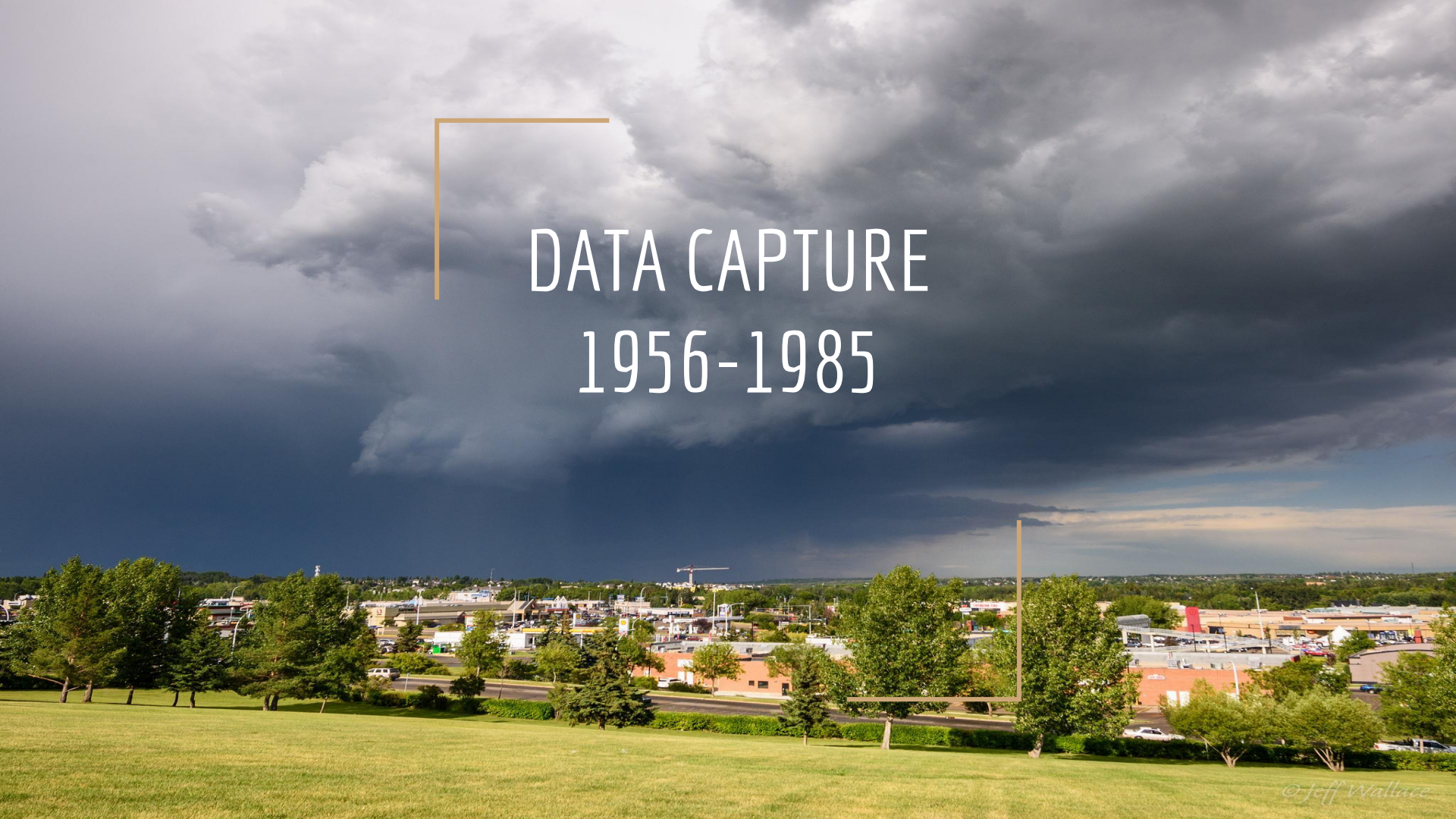

HAIL DATA!
A Data Preservation Saga

Alberta's "Hailstorm Alley"

Sandra Sawchuk, Librarian
Mount Saint Vincent University

Anna Bombak, Data Librarian
University of Alberta

http://mediagestalt.com/presentations/iassist-2018.html


DATA CAPTURE
1956-1985
©Jeff Wallace

COMMUNITY PARTICIPATION

1956 1985

1950 1960 1970 1980 1990 2000 2010 2020

DATA CAPTURE

STORM CHASING VAN

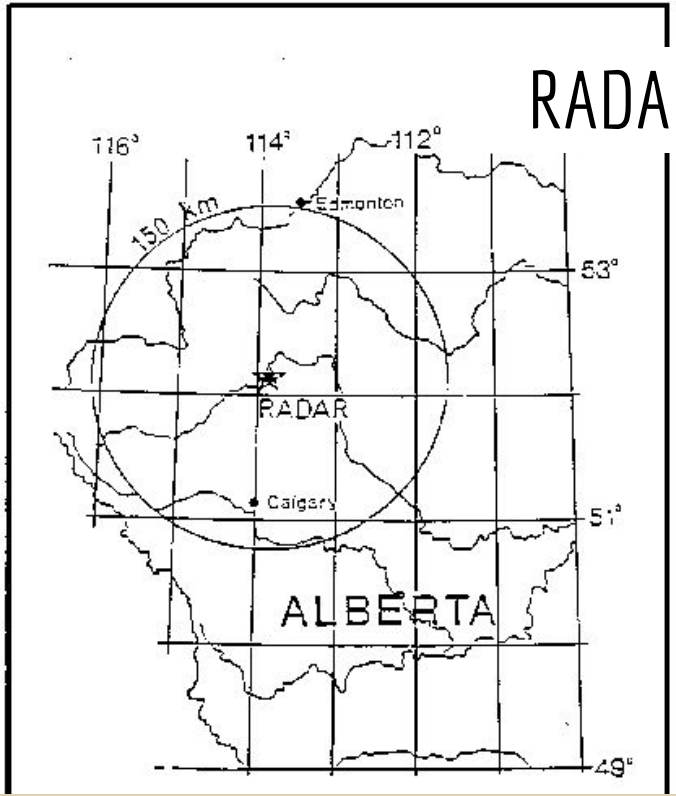
1956

1985

1950 1960 1970 1980 1990 2000 2010 2020

DATA CAPTURE

RADAR



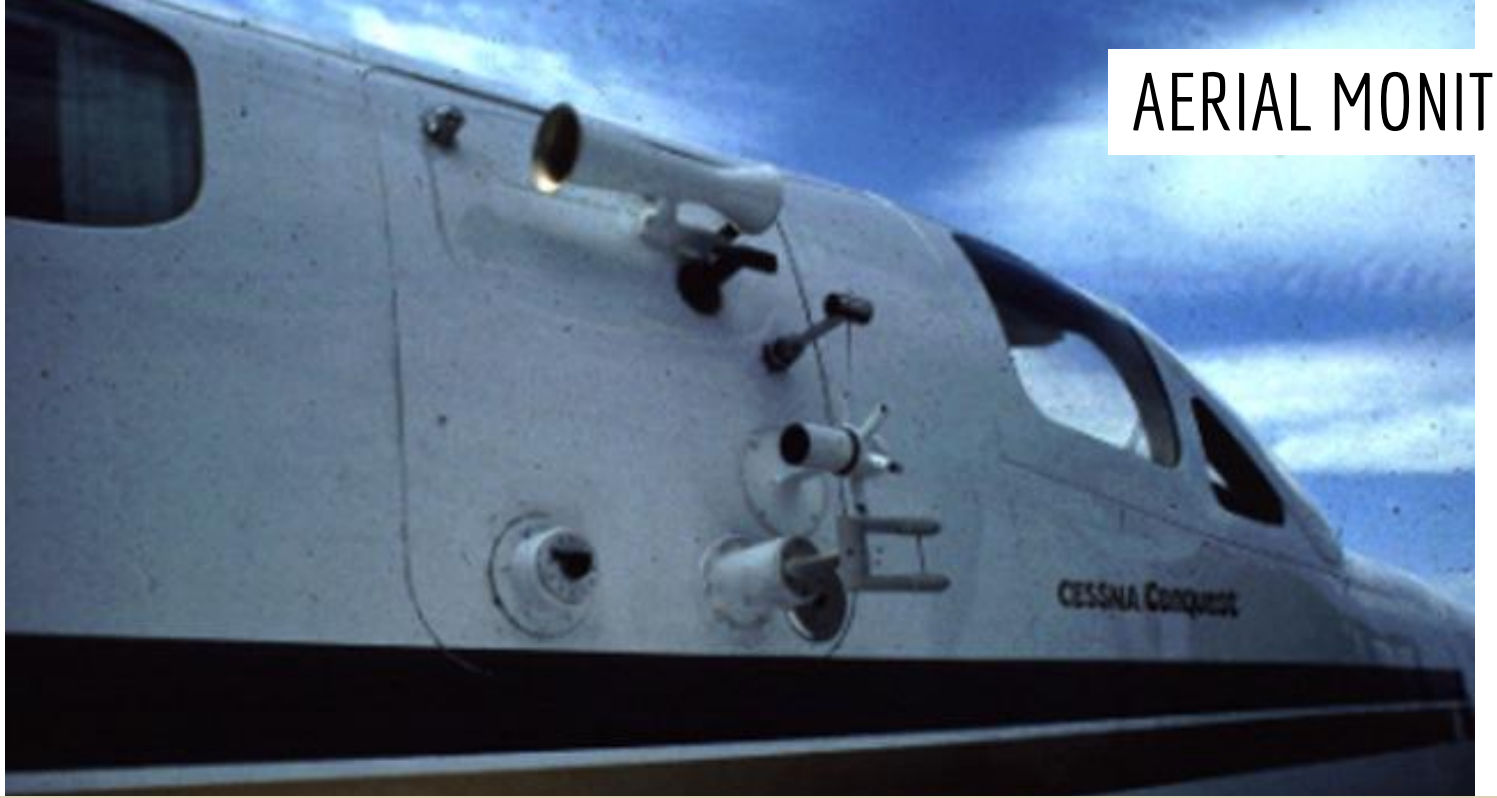
1956

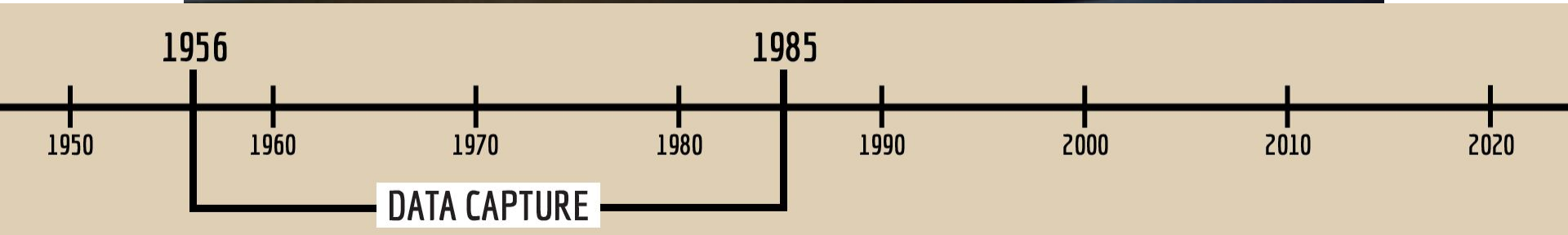
1985

1950 1960 1970 1980 1990 2000 2010 2020

DATA CAPTURE

AERIAL MONITORING





DATA RESCUE & REUSE 1990-1994

MAGNETIC TAPE TO CD-ROM

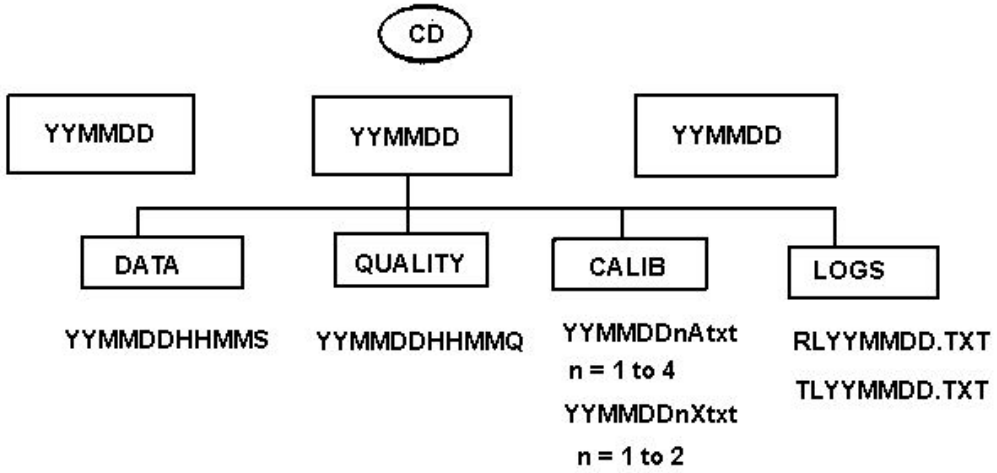
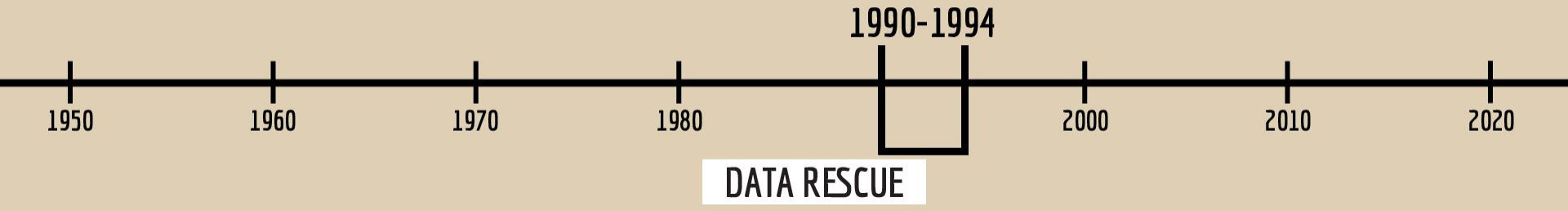


Figure 3: The directory structure of a radar compact disk. Each CD contains a series of daily files (YYMMDD). The radar volume scans, corresponding quailty control reports, calibration text files, and the daily radar and transmitter log files, are found in the DATA, QUALITY, CALIB,and LOGS sub-directories respectively.



9 THESES

a) DEPARTMENT OF METEOROLOGY
McGILL UNIVERSITY, Montreal, Quebec

Balshaw, M., 1967: Behaviour of cumulus turrets in Alberta storms — MSc

b) DEPARTMENT OF GEOGRAPHY (METEOROLOGY)
UNIVERSITY OF ALBERTA, Edmonton, Alberta.

Aktary, N. 1985: Evaluation of ground generator cloud seeding in southern Alberta (1977-1983) — MSc

c) PHYSICS DEPARTMENT
UNIVERSITY OF QUEBEC AT MONTREAL, Montreal, Quebec

Boulet, G., 1985: Utilisation de spectres de tailles de gouttes pour evaluer les performances des radars a — MSc

d) DEPARTMENT OF ATMOSPHERIC SCIENCES
UNIVERSITY OF WYOMING, Laramie, Wyoming

Krauss, T.W., 1981: Precipitation processes in the new growth zone of Alberta hailstorms — PhD

e) DEPARTMENT OF ELECTRONIC SYSTEMS ENGINEERING
UNIVERSITY OF ESSEX, Colchester, United Kingdom

23 PUBLICATIONS

Radar Measurement of Rainfall by Differential Propagation Phase: A Pilot Experiment

M. English, B. Kochtubajda and F.D. Barlow
Resource Technologies, Alberta Research Council
Edmonton, Alberta
and
A.R. Holt and R. McGuinness
Mathematics Department, University of Essex
Wivenhoe Park, Colchester, United Kingdom

[Original manuscript received 13 September 1990; in revised form 12 December 1990]

ABSTRACT *A pilot project concerning the measurement of rainfall by polarization diversity radar, conducted jointly by the Alberta Research Council and the University of Essex in the summer of 1989, is described. The objective of the experiment was to test the theory that differential propagation phase shift can give a better estimate of rainfall with high rain rates (about 30 mm h^{-1} and greater) than reflectivity measurements.*

The project comprised a field experiment that was conducted in central Alberta during the period 20 July to 2 August 1989. The field experiment involved observing storms within about a 70-km radius from Red Deer with the ARC 5-band polarization diversity radar and measuring rainfall at the ground through a network of fixed, volunteer observers and a mobile storm-chase operation.

Theory describing how differential propagation phase may be extracted from the data recorded by the radar system is given.

Data collected on three days during the experiment (27 July, 29 July and 2 August) have been analysed and the results are presented. A total of 31 samples of total rainfall amount were collected on these days. All but three of the 31 radar rainfall amount predictions obtained from the differential propagation phase are within a factor of 2 of the rainfall observed at the ground. In fact, the average ratio of the total rainfall amount predicted from the differential propagation phase to the total rainfall amount measured at the ground is 1.16 for the 31 samples. This suggests that, on average, the total rainfall amount predicted from the differential propagation phase is only 16% higher than that measured at the ground. Of the 31 events, over a third involved some contamination of the differential propagation phase measurement through hail. Furthermore, because the K_{DP} technique does not rely on

4 SCIENTIFIC REPORTS

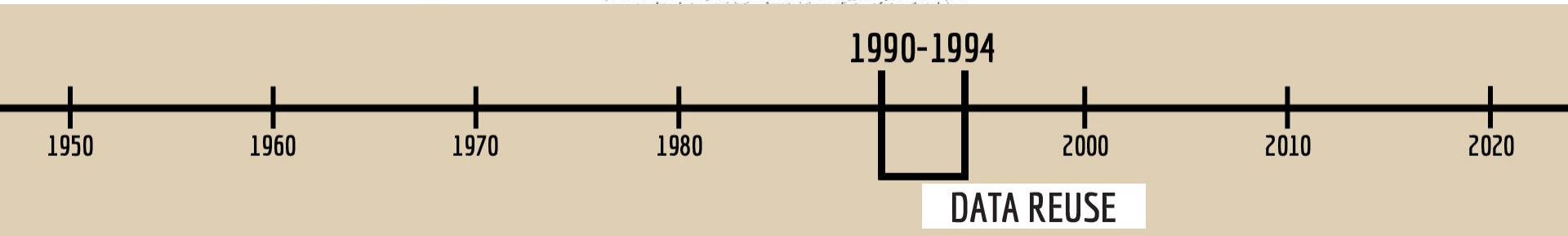
MW-27 Studies of Alberta hailstorms, 1957.
Douglas, R.H., and W. Hitschfeld, May 1958, 79 pp.

MW-29 The motion and erosion of convective storms in severe vertical wind shear.
Hitschfield, W., July 1959, 47 pp.

MW-30 Alberta hail, 1958, and related studies.
Douglas, R.H., Barklie, R.H.D, and N.R. Gokhale, July 1959, 64pp

MW-31 The quantitative display of radar weather patterns on a scale of grey.
Legg, T.H., June 1960. 113 pp.

MW-34 Interim account of hail studies.
Douglas, R.H., J.S. Marshall, and R.H. D. Barklie, Nov, 1960, 46pp.




DATA re-RESCUE
& re-REUSE
2015-2016

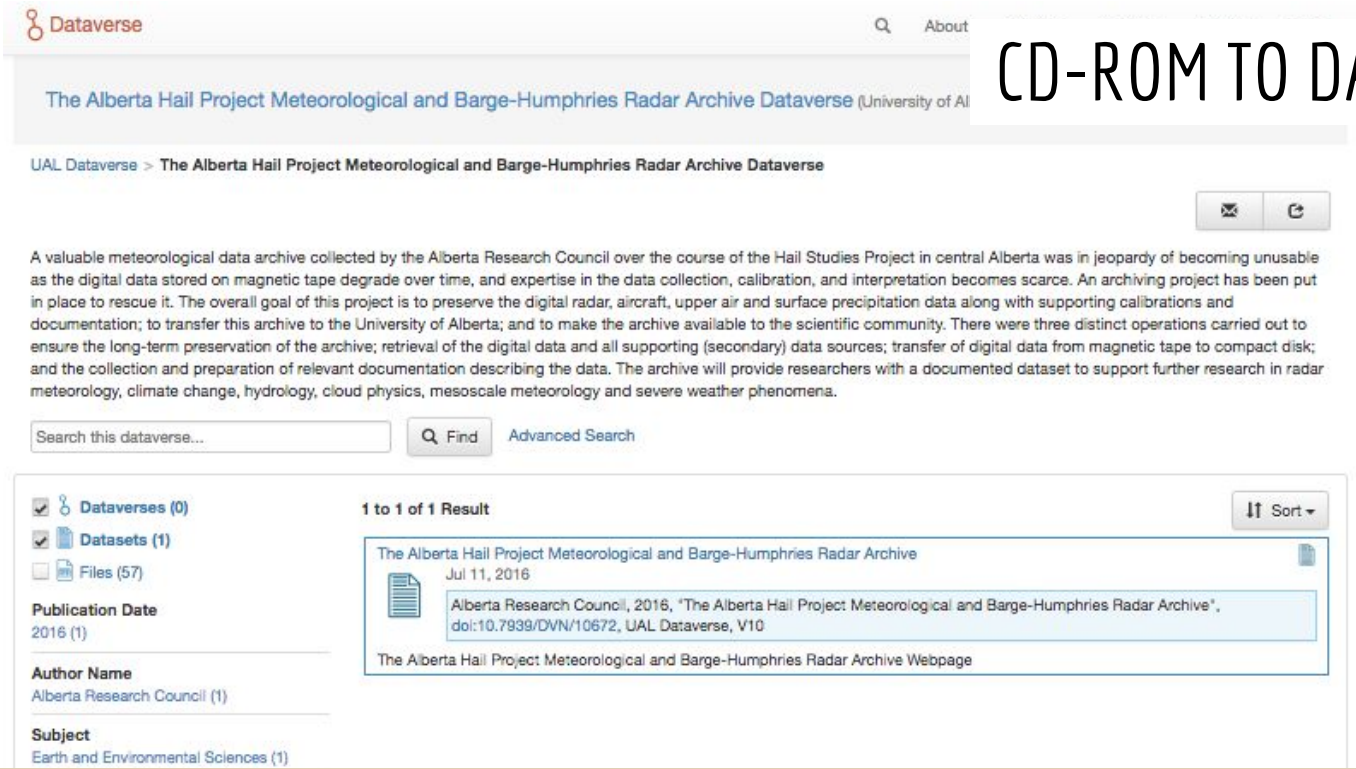
CD-ROM TO DATAVERSE

2015-2016

1950 1960 1970 1980 1990 2000 2010 2020

DATA RE-RESCUE

CD-ROM TO DATAVERSE

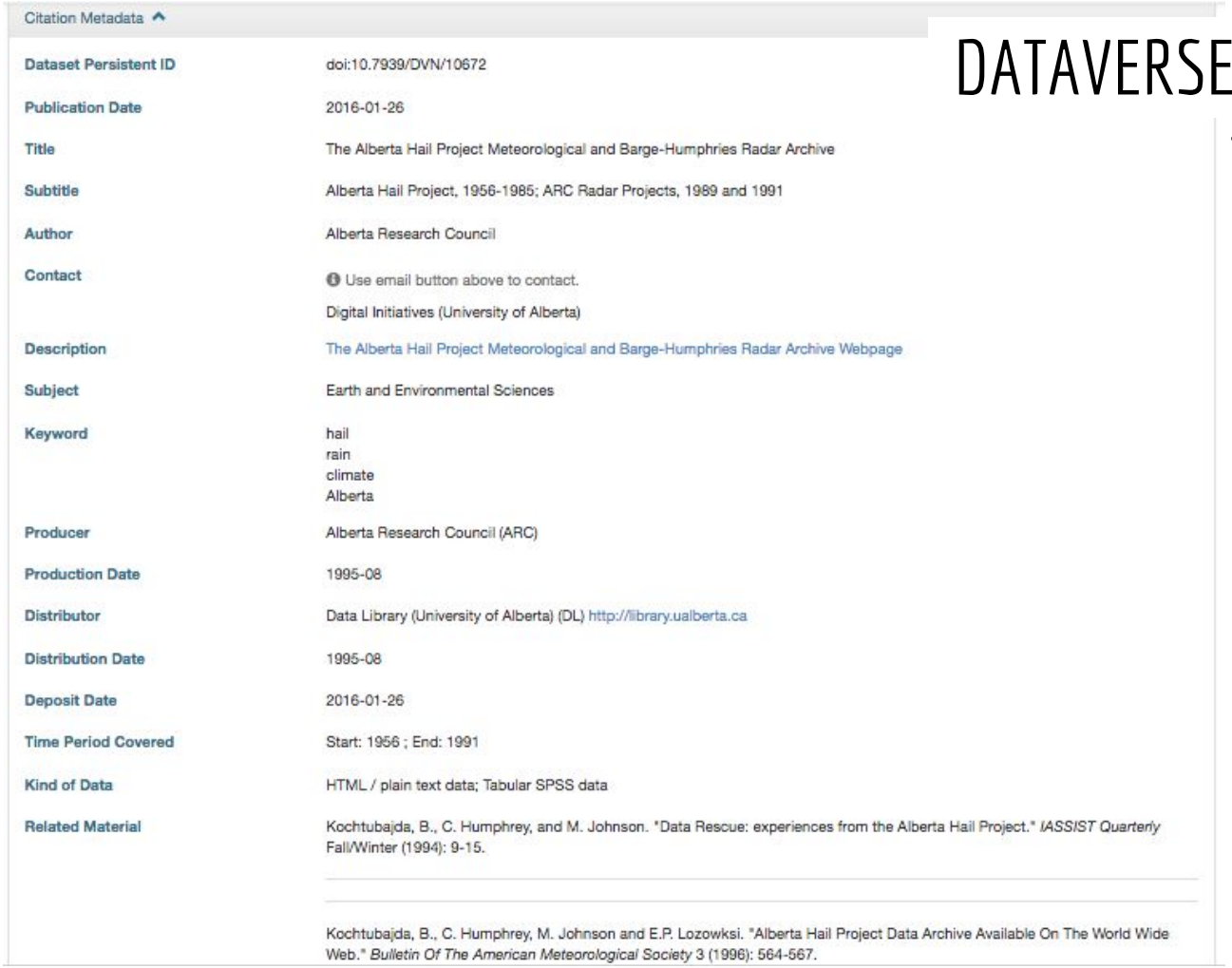


2015-2016

1950 1960 1970 1980 1990 2000 2010 2020

DATA RE-RESCUE

DATAVERSE CITATION METADATA



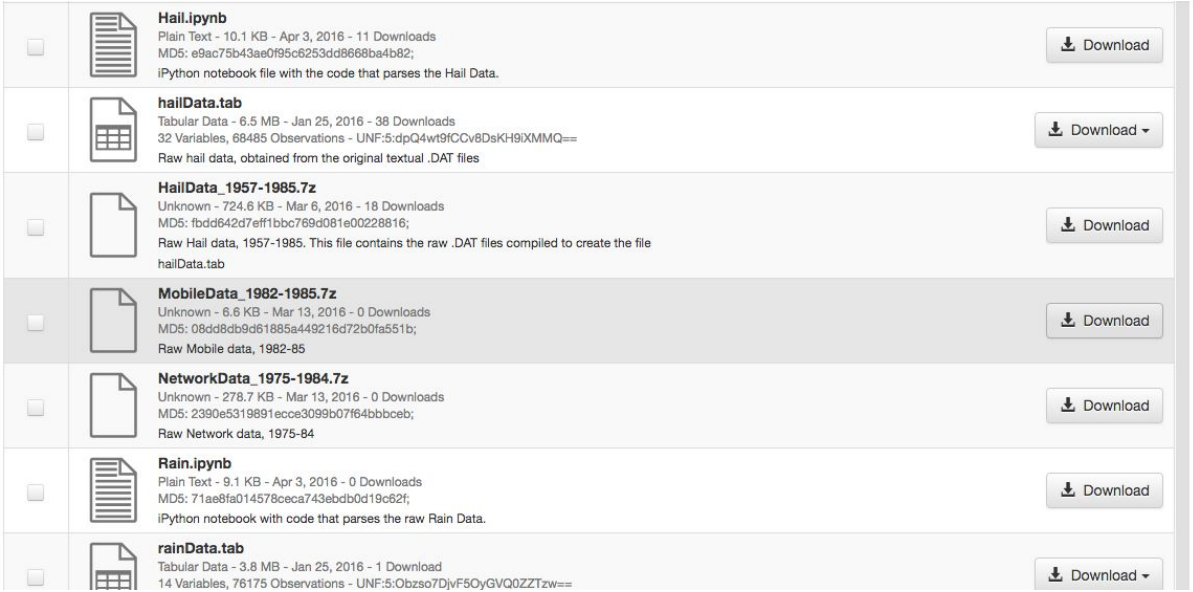
Citation Metadata

Dataset Persistent ID	doi:10.7939/DVN/10672
Publication Date	2016-01-26
Title	The Alberta Hail Project Meteorological and Barge-Humphries Radar Archive
Subtitle	Alberta Hail Project, 1956-1985; ARC Radar Projects, 1989 and 1991
Author	Alberta Research Council
Contact	Use email button above to contact. Digital Initiatives (University of Alberta)
Description	The Alberta Hail Project Meteorological and Barge-Humphries Radar Archive Webpage
Subject	Earth and Environmental Sciences
Keyword	hail rain climate Alberta
Producer	Alberta Research Council (ARC)
Production Date	1995-08
Distributor	Data Library (University of Alberta) (DL) http://library.ualberta.ca
Distribution Date	1995-08
Deposit Date	2016-01-26
Time Period Covered	Start: 1956 ; End: 1991
Kind of Data	HTML / plain text data; Tabular SPSS data
Related Material	Kochtubajda, B., C. Humphrey, and M. Johnson. "Data Rescue: experiences from the Alberta Hail Project." *IASSIST Quarterly* Fall/Winter (1994): 9-15. Kochtubajda, B., C. Humphrey, M. Johnson and E.P. Lozowksi. "Alberta Hail Project Data Archive Available On The World Wide Web." *Bulletin Of The American Meteorological Society* 3 (1996): 564-567.

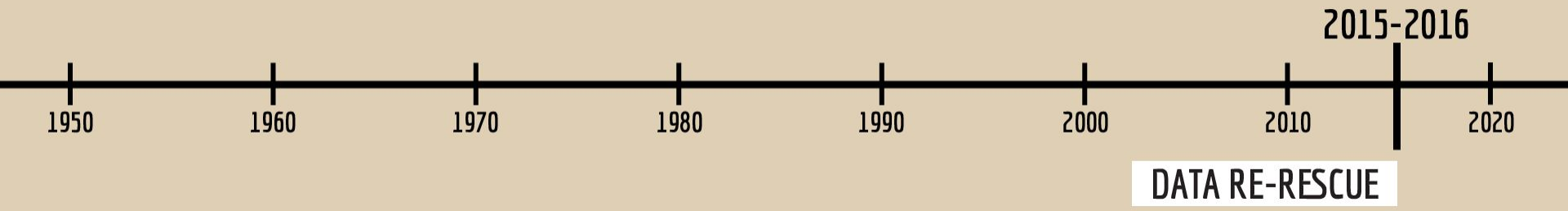
This information can be harvested and made discoverable by other repositories:

- <odesi>
- Federated Research Data Repository (FRDR)
- Education Research Archive (University of Alberta)
- other Dataverses

DATAVERSE FILE DOWNLOAD



All the files and their supporting documentation can be downloaded. This archive contains all of the original data, the HTML directory of the original data rescue project webpage from the 1990s, and the Jupyter Notebook files used to parse the data in 2016.

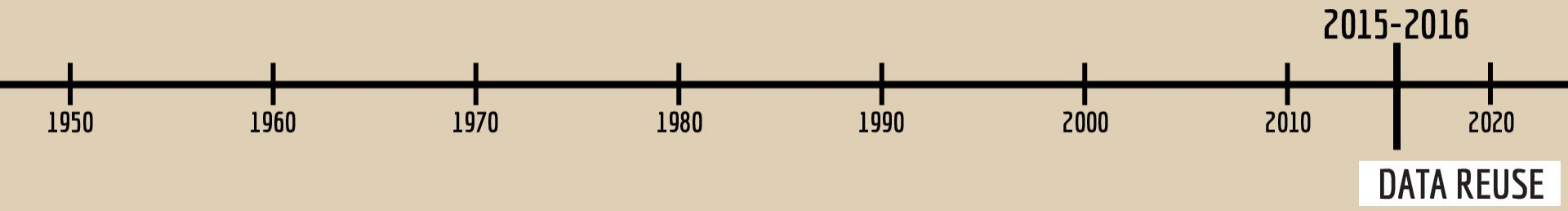


```
1576 4   11938 351800 109999999999939999999999999999199999
1576 4   1 939 551630  799999999999299999999999999999999999
1576 4   133292241335  299999999999299999999999199999199999
1576 4110014202641700 159999999999949999999999999999199999
1576 5   13619 151755 209999999999939999999999999999199999
2576 5   1 320 151740 209999999999939999999999999999199999
2576 5 1001020 151740 209999999999939999999999999999199999
1576 5  102825 151630 129999999999949999993  76999999199999
1576 5   11227 351530 309999999999929999999999999999099999
1576 510003327 451600 259999999999939999999999999999199999
1576 5 1002238 451740 109999999999939999999999999999099999
1576 5111114192441830 189999999999949999999999999999199999
1576 5   1 1192541745 109999999999929999999999999999185999
1576 5   121192541810 259999999999939999999999999999099999
1576 5100032192541815 109999999999939999999999999999099999
1576 5  1021202541730 259999999999939999999999999999099999
1576 5   128212541800 209999999999939999999999999999199999
1576 5  10 6222541600 159999999999939999999999999999199999
1576 5  1025192641800 259999999999939999999999999999099999
1576 5 10034202641730 1599999999 254999999999999999999085999
1576 5110136212641600 159999999999939999999999999999199999
1576 51111 1222641600 159999999999939999999999999999199999
1576 5100011222641830 109999999999939999999999999999199999
1576 5111114222641700 209999999999939999999999999999198999
```

HAIL CARD RAW DATA

Each line represents a hail card that was mailed in or reported by phone.

Each number, or group of numbers, represents a unique data point: including the date and time, location, size of hail, and the nature of crop damage.



HAIL CARD CODEBOOK

Time & Date

Geographical Data

(use 1-29 to code rain cards)
[ ] - represents columns

[1] Blank
[2] TYPE OF REPORT
1 Mailed in
2 Tele in
3 Network rpts.
4 Tele Survey area
6 Rain only
7 No hail/rain
8 Out of proj.

[3] YEAR
[4] (Last two digits of year)

[5] MONTH
4 April
5 May
6 June
7 July
8 August
9 September
0 October

[6] DAY OF MONTH
[7] (e.g. 03,14, etc.)

[8] QUARTER (8,9,10,11)
[9] (NE,SE,SW,NW)
[10] 0 - No
1 - Yes
[11] 9999 Missing

[12] SECTION
[13] 01 - 36

[14] TOWNSHIP
[15] 01 - 99

[16] RANGE
[17] 01 - 30

[18] MERIDIAN
4,5 or 6

[19] TIME OF HAIL - D.S.T.
[20] 0000-2359

[39]
(Largest Stones)
1 Spacing
2 Just covered
3 Depth
9 Missing or Melted

[40] ACTUAL SPACING OR DEPTH
Spacing in mm or depth
[41] 9998 if greater than
9999mm
2.5mm-0003
25.4mm-0025
[42] 2544mm-0254
[43] 2544.4mm-2544
00-Just Covered

[44] HAIL SPACING OR DEPTH
1 Spacing
2 Just covered
3 Depth
9 Missing or Melted

[45] ACTUAL SPACING OR DEPTH
Spacing in mm or depth
in mm.

[46]
[47] 9998 if 9999 MM or more
[48] 00 - Just Covered

[49] SOFT HAIL
0-No 1-Yes 8-Missing

% of soft hail
99 Missing or hard hail
[50] 01-98% of soft stones
[51] 00-100% soft stones

[23] DURATION - Minutes
[24] 000-998
[25] 999 - missing
1/2 min. - 001

[26] TIME OF RAIN - D.S.T.
[27]
[28] Same as hail
[29] 8888 - No Rain

[30] DURATION - MINUTES
[31]
[32] Same as Hail
999 Missing or No Rain

[33] MEASURED ESTIMATED MISSING
0 1 9

[34] RAINFALL EXAMPLE
[35] 000 - Trace 12.2-122
[36] 001 - 998 mm 1.2-012
998 - Not mentioned .2-002

Hail Size

HAIL SIZE
1 shot
2 pea
3 grape
4 walnut
5 golfball
6 larger
9 missing
[37] Maximum size
[38] Common Size

3 egg 9 missing

[53] SURFACE TEXTURE
1 smooth 8 others
2 raspberry 9 missing
3 knobby

[54] WINDS
0 light (0 - 15 km/h)
1 moderate (15 - 40 km/h)
2 strong (40 - 65 km/h)
3 severe (over 65 km/h)
9 missing or gusty

[55] LIGHTNING
0-No 1-Yes 9-Don't know

[56] DAMAGE PERCENTAGE
00 - 97
[57] 98 code 100%
99 missing

Crop Type & Damage

[58] CROP TYPES
0 mixed grain 5 rapeseed
1 wheat 6 garden
2 barley 7 rye
3 oats 8 other
4 hay 9 missing

[59] GROWTH STAGE
1 3-5 leaf 5 dough
2 shot blade 6 mature
3 heading 9 missing
4 blooming

[60] STORM APROACH DIRECTION
1-N 4-SE 7-W
2-NE 5-S 8-NW
3-E 6-SW 9-MISSING

PARSING THE DATA IN PYTHON

Reading the .DAT files...

```
Hail/76HAIL.txt
 176525100029382441010015100018002542299999200001339909992l9
 176525000135320551110015111401111272999999200001994109004l9
 176527001003452349999010999912000762299999100250994129009999
 176529100020490641645012165701099992299999920000099411900919
 176529001010381541200005115502599993299999920000125411900999
 176531010020460451637009162505599992999999101520994109000019
 176531000132370751720030171004599993299999300201404319993999
```

Regular expression in Python parses each line of text...

```
r'^(\s)(\d{1})(\d{2})(\d{1})(\d{2})(\d{4})(\d{2})(\d{2})(\d{2})(\d{1}
)(\d{4})(\d{3})(\d{4})(\d{3})(\d{1})(\d{3})(\d{1})(\d{1})(\d{1})(\d{4
})(\d{1})(\d{4})(\d{1})(\d{2})(\d{1})(\d{1})(\d{1})(\d{1})(\d{2})(\d{
1})(\d{1})(\d{1})
```

...storing the numbers as a comma separated list

```
[(' ', '1', '76', '5', '25', '1000', '29', '38', '24', '4', '1010',
'015', '1000', '180', '0', '254', '2', '2', '9', '9999', '2', '0000',
'1', '33', '9', '9', '0', '9', '99', '2', '1', '9')]
```

```
filenames = []
for files in list_text
    files = get_filena
    filenames.append(
```

```
docs = []
for filename in list_t
    docs.append(read_
```

```
data_2 = []
for doc in docs:
    data_2.append(re.
```

```
alldata = [line for su
```

```
with open('allHail.csv
    w = csv.writer(f)
    w.writerows(alldat
```

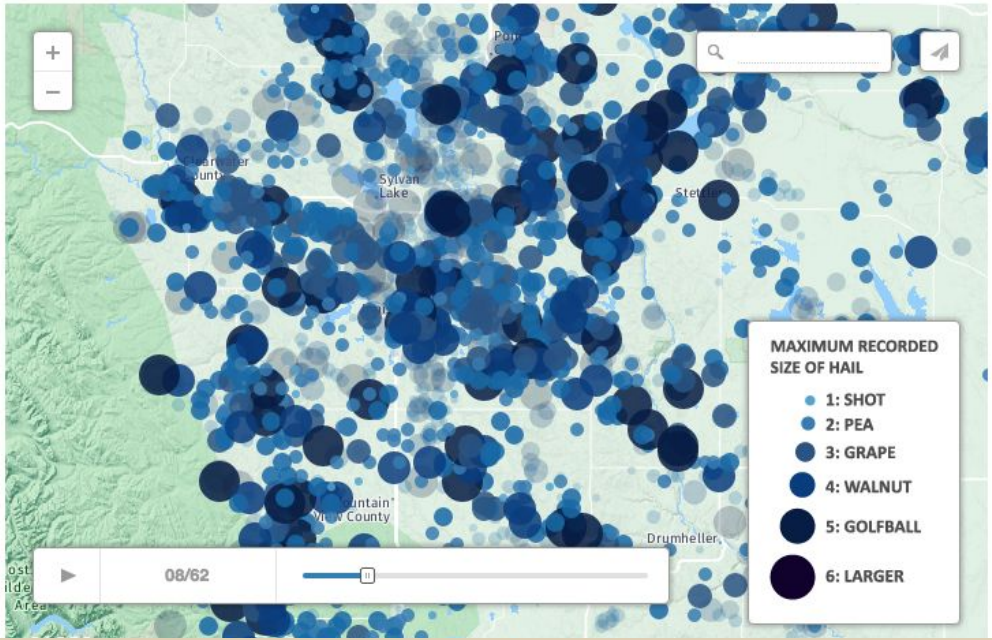
ASSIGNING THE LATITUDE/LONGITUDE

MER	RGE	TWP	SEC	QTR	UID	POINT_X	POINT_Y
4	26	54	1	NE	42605401NE	-113.69518766800	53.63978045280
4	26	54	1	NW	42605401NW	-113.70731870800	53.63979984040
4	26	54	1	SE	42605401SE	-113.69523560000	53.63254647670
4	26	54	1	SW	42605401SW	-113.70736036000	53.63255919770
4	26	54	2	NE	42605402NE	-113.71975873000	53.63981148820
4	26	54	2	NW	42605402NW	-113.73191530700	53.63982313720
4	26	54	2	SE	42605402SE	-113.71978815300	53.63257969450
4	26	54	2	SW	42605402SW	-113.73192437300	53.63259369280
4	26	54	3	NE	42605403NE	-113.74432152200	53.63983505360
4	26	54	3	NW	42605403NW	-113.75633853300	53.63986591410
4	26	54	3	SE	42605403SE	-113.74435220500	53.63260311790
4	26	54	3	SW	42605403SW	-113.75643163200	53.63261026020
4	26	54	4	NE	42605404NE	-113.76879373800	53.63987159520
4	26	54	4	NW	42605404NW	-113.78107732300	53.63987288620
4	26	54	4	SE	42605404SE	-113.76889360300	53.63263620130
4	26	54	4	SW	42605404SW	-113.78110987500	53.63263092930
4	26	54	5	NE	42605405NE	-113.79358362000	53.63985831810
4	26	54	5	NW	42605405NW	-113.80573133000	53.63984130250
4	26	54	5	SE	42605405SE	-113.79358442600	53.63262176060
4	26	54	5	SW	42605405SW	-113.80573734000	53.63261552920
4	26	54	6	NE	42605406NE	-113.81819508700	53.63985002870
4	26	54	6	NW	42605406NW	-113.83017788200	53.63987464750
4	26	54	6	SE	42605406SE	-113.81821128000	53.63263257650
4	26	54	6	SW	42605406SW	-113.83017845500	53.63265170930
4	26	54	7	NE	42605407NE	-113.81818000100	53.65430799670
4	26	54	7	NW	42605407NW	-113.83017158100	53.65432225580
4	26	54	7	SE	42605407SE	-113.81818924600	53.64707858120
4	26	54	7	SW	42605407SW	-113.83017558000	53.64709823400
4	26	54	8	NE	42605408NE	-113.79357587600	53.65431791870
4	26	54	8	NW	42605408NW	-113.80571841300	53.65430524720
4	26	54	8	SE	42605408SE	-113.79358074400	53.64708617600

The Alberta Township Survey System is a grid network dividing the province into equal-sized parcels of land. It has been in use since the early 1900s.

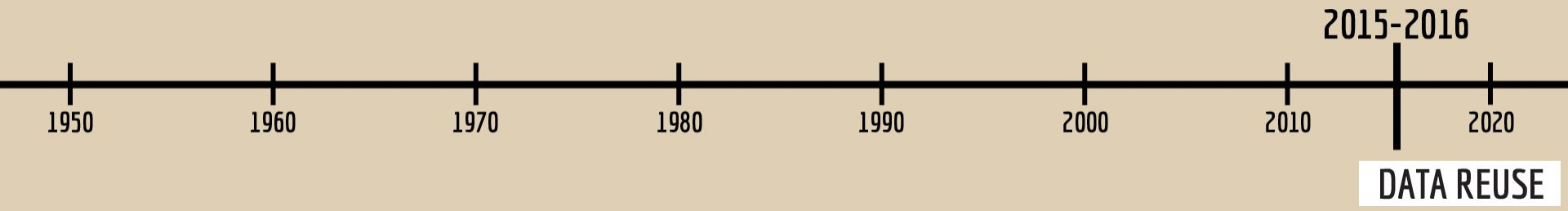
The unique codes from the hail cards were matched to their corresponding latitude and longitude points in order to create the map in CartoDB.

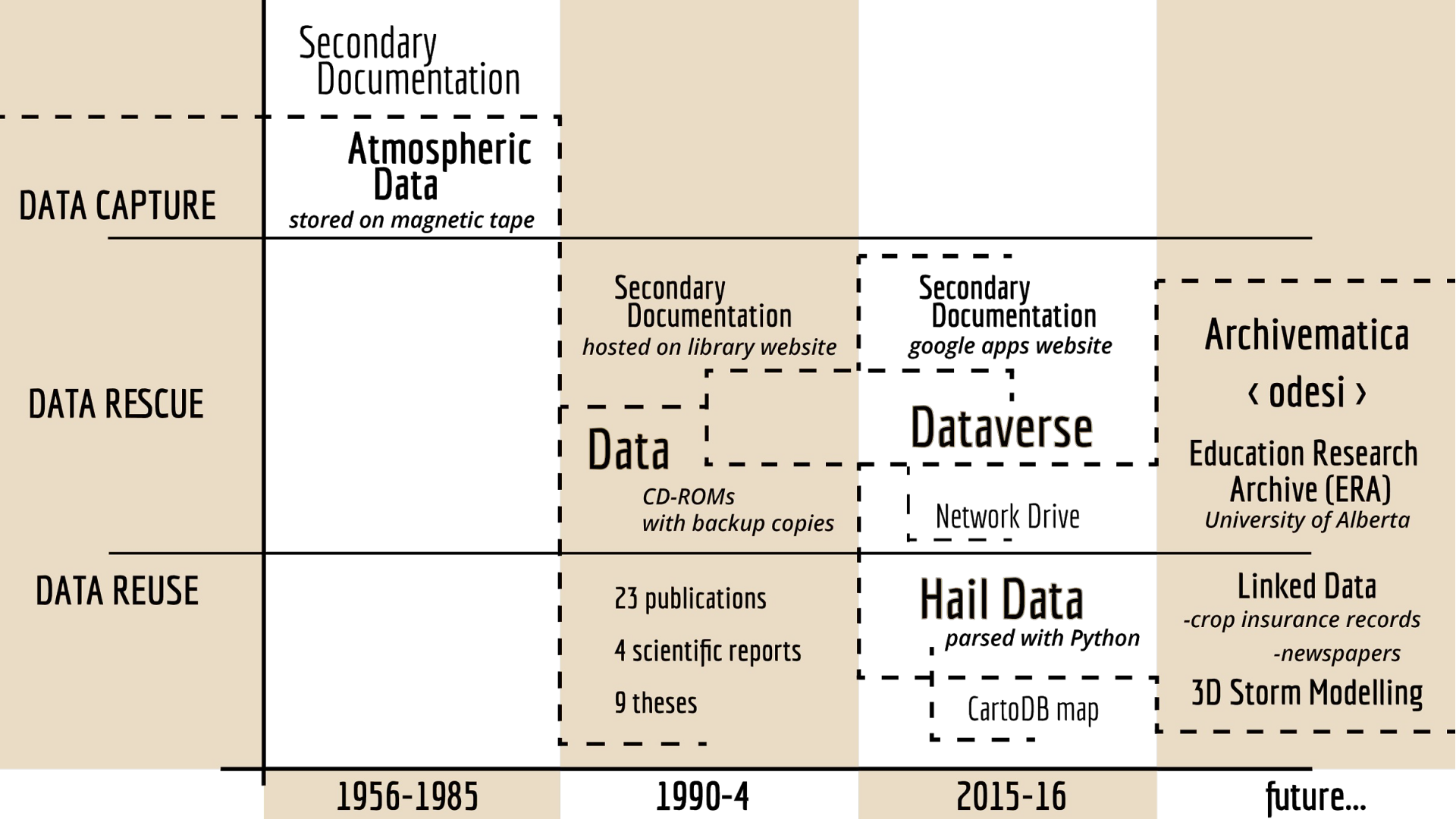
MAPPING HAIL SIZE AND FREQUENCY



The cleaned dataset (with the GIS coordinates) was imported into CartoDB, a free, online mapping program.

The maps can be completely customized using CSS and HTML.




Secondary Documentation
DATA CAPTURE
Atmospheric Data
stored on magnetic tape
DATA RESCUE
Secondary Documentation
hosted on library website
Secondary Documentation
google apps website
Archivematica
‹ odesi ›
Dataverse
Data
CD-ROMs
with backup copies
Network Drive
Education Research Archive (ERA)
University of Alberta
DATA REUSE
23 publications
4 scientific reports
9 theses
Hail Data
parsed with Python
Linked Data
-crop insurance records
-newspapers
CartoDB map
3D Storm Modelling
1956-1985
1990-4
2015-16
future...

THE END ?

anna.bombak@ualberta.ca

sandra.sawchuk@msvu.ca
twitter/github @mediagestalt
website mediagestalt.com