OPEN RADAR
SCIENCE FOR FUN AND, YES, EVEN PROFIT.

SCOTT COLLIS
Atmospheric Scientist, Argonne National Laboratory.
Senior Institute Fellow, Northwestern University.

ROBERT JACKSON, ZACH SHERMAN AND MARK PICEL
Argonne National Laboratory

AND SO MANY MORE....
THIS TALK

What started as a way of building a radar retrievals/QC pipeline turned into a popular community radar toolkit

- This talk will:
  - Give a bit of history as to how the Python ARM Toolkit came into being.
  - Introduce Py-ART for those that do not know what it is.
  - Show some example usage.
  - Sustainable growth: The Py-ART roadmap.
  - Introduce the Open Radar Partnership.
  - Some science we have done using Py-ART and open weather radar data.
THIS TALK

What started as a way of building a radar retrievals/QC pipeline turned into a popular community radar toolkit

- This talk will:
  - Give a bit of history as to how the Python ARM Toolkit came into being.
  - Introduce Py-ART for those that do not know what it is.
  - Show some example usage.
  - Sustainable growth: The Py-ART roadmap.
  - Introduce the Open Radar Partnership.
  - Some science we have done using Py-ART and open weather radar data.
A LITTLE HISTORY

Building a retrievals and quality control framework across a diverse network of radars

- Our journey into open source software did not start "properly". It was not scoped, designed, frameworked, engineered and tested initially.

- ARM has a network of 20+ radars. Made by four vendors. Each vendor has their own format.

- Py-ART started as a desire to consistently represent radar data in Python.
RADAR DATA
The shape of it....

- Weather radars transmit a pulse of energy (in GHz range, pw ~ 10’s of meters) and gate the receiver to achieve ranging.
- The basic data form are rays. Collection of data collected at different (timed) range.
- For most radars the number of gates remain constant so data can be stored in a number of ND-Arrays of (time,range).
- In addition, for steerable (scanning) radars information about the pointing of the antenna must be stored.
Philosophy: It’s all about the data model.

- Py-ART’s central core is a data model for gated data with pointing information.
- Py-ART created a way of representing radar data in the Python programming language that mirrors the CF-Radial standard.
- Py-ART has a cloud functions to correct, retrieve and grid radar data.
- By keeping a limited scope Py-ART aims to “do less better”.
- There is now a rich ecosystem of packages that interact: ART-View, CSU tools, PyTDA, Multidop, TINT, Py-Hail, Py-DDA, Unravel... to name a few.

Animation courtesy of users Marcus van Lier-Walqui and Sara E. Lytle
Data: Andrei Lindenmaeir – ARM Mentor
TYPICAL USAGE OF PY-ART

Notebooks – Single file application

Scale to a cluster

- Jupyter notebooks give you the ability to interact in a tight feedback cycle with radar data.
- Great for adjusting parameters like membership functions or weights.
- Get that right for a single file and then port (using your favorite IDE, PyCharm!) to an executable.
- Many of us have access to clusters or at least multi-core machines. Tools like IPyCluster and Dask allow single “granules” to be mapped to many workers.
- \( \sim 20M \) voxels/grid = \( \sim 300M \) voxels/s inc. reading wsr88d and writing CF-Radial.
WHERE'S THE KABOOM?!

THERE'S SUPPOSED TO BE AN EARTH-SHATTERING KABOOM!
PY-ART IS GLOBAL
Self reported installations of Py-ART
The power of community software. But, you need standards, unit testing and continuous integration.

Only scollis, jjhelmus, zssherman, rcjackson and rumpkie have received ARM funds. The other 28 have not...
Making radar codes open source is not free. Non-ARM funded people put in funded and spare time and ARM funded folks have a line item for upkeep.

We have many automated tools to check when things break. But things still need fixing. Although we always work to minimize critical failure paths.

To this end we need to ensure funds spent benefit ARM and its stakeholders. We developed the Py-ART roadmap to do this.

https://commons.wikimedia.org/wiki/File:Isummit_2008,_Japan,_free_beer.jpg
OPEN RADAR PARTNERSHIP
We are all in this together…

- Python efforts in Radar Meteorology nucleated in parallel in Europe and the USA.
- In addition the long standing TITAN software cloud received NSF support for a modernization surge.
- A short course at ERAD 2014 led to a co-written paper which lead to a rotating course between ERAD and the AMS radar conference.
- We now have a web presence and a growing ecosystem of identified packages.
PROFIT: DOE PROCESSING

All done in Parallel Dask or Py-Spark

- Interpolated Sonde or Sounding
- a1/b1 data
- Clutter
- Vr Texture
- Temp. at gate
- Pseudo SNR
- Gate ID
- Dealiased Vr
- Processed PhiDP
- Processed KDP
- Specific Attenuation
- Rainfall Rate (A)

1-3 Months of staged data

Clutter detection

Texture and pseudo SNR calculated

Append Temp @ gate

Fuzzy Logic Gate ID

Region Based Dealias

LP PhiDP

Sobel KDP

Spider Removal

Z-Phi Specific Attenuation

Spec. Atten Rainfall

Visualization and Summary

Databases, quick looks etc...

1-3 Months of staged CMAC2.0
PROFIT: TRACER

How many cells will we see with a research radar in Houston?

- Could we propose to ARM to deploy, along with the rest of the ARM mobile facility, the C-SAPR2 deployable Dual-Pol research radar?
- And can we ask to receive engineering support to adaptively follow storm cells based on what is seen on the radar or nearby NEXRAD KHGX?
- Before we do any planning or theorizing on how we would operate we first need to understand Houston convection.
- When (seasonally) are we most likely to get nice isolated cells?
- What is the behavior of these cells?
  - Life cycle
  - Formation points
  - Dissipation points
TINT: TINT Is Not TITAN

- TINT is cell tracking made simpler. Takes a series of Py-ART grids, returns Pandas data frames -> CSV.
- Two step: Image cross correlation gives bulk motion. Search picks closest cell to back propagation.
- Also detects if cell is isolated.

https://github.com/openradar/TINT
FOUR YEAR CLIMATOLOGY OF CELLS OVER HOUSTON

Open data + open source software = awesome

- Use Dask to map-reduce. ~512 cores, took O(2days)
- 7TB NEXRAD data -> 100MB of track data.
- Question being addressed: If we sent a radar to Houston how many storms would we see with the full lifecycle in the MUR?
- Also: What time of year should we deploy if we have limited resources (which of course we always have). If we were to have an IOP for cell tracking when is best?

![Graph showing cell count over time and distance]

- 411 @ 70 km
- 2442 @ 150 km
- 4843 @ 200 km

![Bar graph showing cell count by month]

- Month 6 and 7 have the highest number of cells.
FOUR YEAR CLIMATOLOGY OF CELLS OVER HOUSTON

Open data + open source software = awesome

- Use Dask to map-reduce. ~512 cores, took O(2days)
- 7TB NEXRAD data -> 100MB of track data.
- Question being addressed: If we sent a radar to Houston how many storms would we see with the full lifecycle in the MUR?
- Also: What time of year should we deploy if we have limited resources (which of course we always have). If we were to have an IOP for cell tracking when is best?
PROFIT… INTO THE FUTURE

Just funded and future open science projects

- **SAVUER**: 1M Project between Northwestern and Argonne looking at urban hydrology and socioeconomic impact of high gradient events.

- Grid modernization: All weather radar data in the USA is free at **real time** and is very easy to obtain from AWS. We are working with DoE office of electricity building tools for integrating weather data into a smart grid.


- Expanding beyond radar…..
RESEARCH TO OPERATIONS
There is one thing we can all agree on: We all want better software
RESEARCH TO OPERATIONS

- Community software has a very important role in R2O:
  - Democratizing access to radar data.
    Allowing communities to use radar data who previously did not have access.
  - Providing examples to be built on.
    Standing on the shoulders of giants.
  - Effective communication of science
    SHOW ME THE CODE!
  - Acting as a net gathering diverse knowledge.

- **Open data rapidly accelerates all of the above. Australia opening radar data means development can occur on AUSTRALIAN data and not just NEXRAD et al.**
RESEARCH TO OPERATIONS

- Community software has a very important role in R2O:
  - Democratizing access to radar data. Allowing comminutes to use radar data who previously did not have access.
  - Providing examples to be built on. Standing on the shoulders of giants.
  - Effective communication of science. SHOW ME THE CODE!
  - Acting as a net gathering diverse knowledge.

- **Open data rapidly accelerates all of the above. Australia opening radar data means development can occur on AUSTRALIAN data and not just NEXRAD et al.**

Precipitation Nowcasting: Leveraging Deep Recurrent Convolutional Neural Networks

Alexander Heye*, Karthik Venkatesan†, Jericho Cain†

Cray, Inc.
Seattle, Washington

*aheye@cray.com
†kvenkatesa@cray.com
†jecain@cray.com
RESEARCH TO OPERATIONS

- Community software has a very important role in R2O:
  - Democratizing access to radar data.
    Allowing communities to use radar data who previously did not have access.
  - Providing examples to be built on.
    Standing on the shoulders of giants.
  - Effective communication of science.
    SHOW ME THE CODE!
  - Acting as a net gathering diverse knowledge.

- Open data rapidly accelerates all of the above. Australia opening radar data means development can occur on AUSTRALIAN data and not just NEXRAD et al.
COME ON IN!

- Official source code repository: https://github.com/ARM-DOE/pyart
- HTML documentation: http://arm-doe.github.io/pyart-docs-travis/
- Examples: http://arm-doe.github.io/pyart/dev/auto_examples/index.html
- Mailing List: http://groups.google.com/group/pyart-users/
- Issue Tracker: https://github.com/ARM-DOE/pyart/issues
- Facebook: https://www.facebook.com/PythonRadar/
- Twitter: https://twitter.com/py_art?lang=en

OPEN RADAR PARTNERSHIP:
HTTP://OPENRADARSCIENCE.ORG/

THANK YOU FOR YOUR TIME
SCOLLIS@ANL.GOV

This presentation has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. This research was supported by the Climate Model Development and Validation activity funded by the Office of Biological and Environmental Research in the US Department of Energy Office of Science. Computing resources were provided by the Laboratory Computing Resource Center of Argonne National Laboratory.