Adapting to the Atmosphere, Atmospheric characterisation and AO conference, Durham, 15th-17th Sept. 2014, www.dur.ac.uk/cfai/sitecharacterisation/

On-Sky Validation Of An Artificial Neural Network Tomographic Reconstructor With The Canary AO System

James Osborn, D. Guzman, F.J. de Cos Juez, A.G. Basden, T.J. Morris, E. Gendron, T. Butterley, R.M. Myers, A. Guesalaga, F. Sanchez Lasheras, M. Gomez Victoria, M.L. Sanchez Rodriguez, D. Gratadour and G. Rousset

Artificial Neural Networks

- A neural network is a parallel computational architecture composed of individual elements or 'neurons'.
- A technique to build a connection map between given inputs and desired outputs
- Each neuron has a weighting function (usually hyperbolic tangent sigmoid)
  - A linear weighting function would develop an ANN identical to a standard matrix vector multiplication
  - A non-linear activation function is used to represent complicated non-linear datasets

Artificial Neural Networks and Adaptive Optics

- Focal plane wavefront sensing (Angel et al., Nature 1990, Sandler et al. Nature 1991 and Lloyd-Hart et al. ApJL, 1992)
- SH centroiding – inconclusive (Montera et al. Appl. Opt., 1996)
- Temporal prediction of modal wavefront (up to m=6) from *a priori* off-axis measurements (Weddell and Webb, Hybrid Intelligent Systems, 2006)

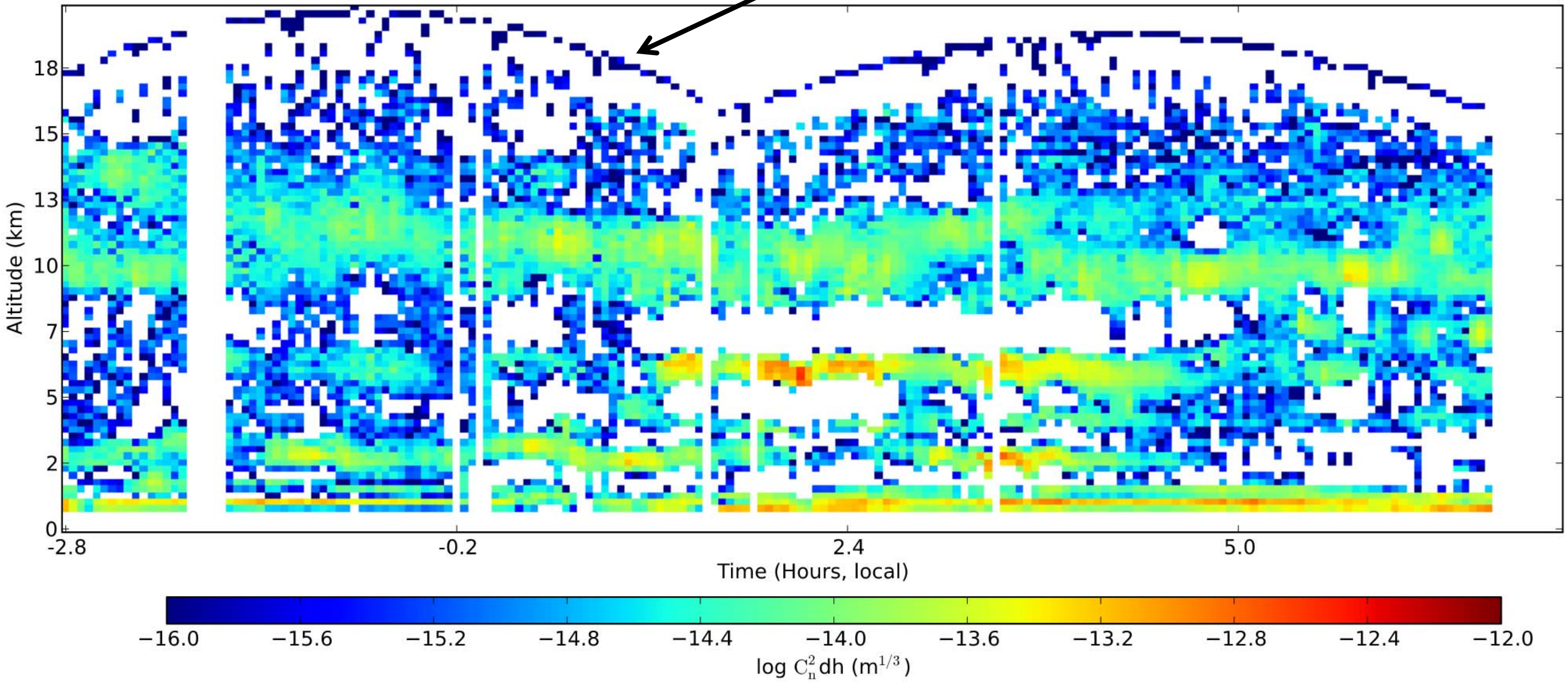
Motivation

- Build a ANN tomographic reconstructor which is insensitive to changing atmospheric conditions.

Results from Stereo-SCIDAR

13/09/2013, JKT 1m, La Palma,

Artefact from covariance function



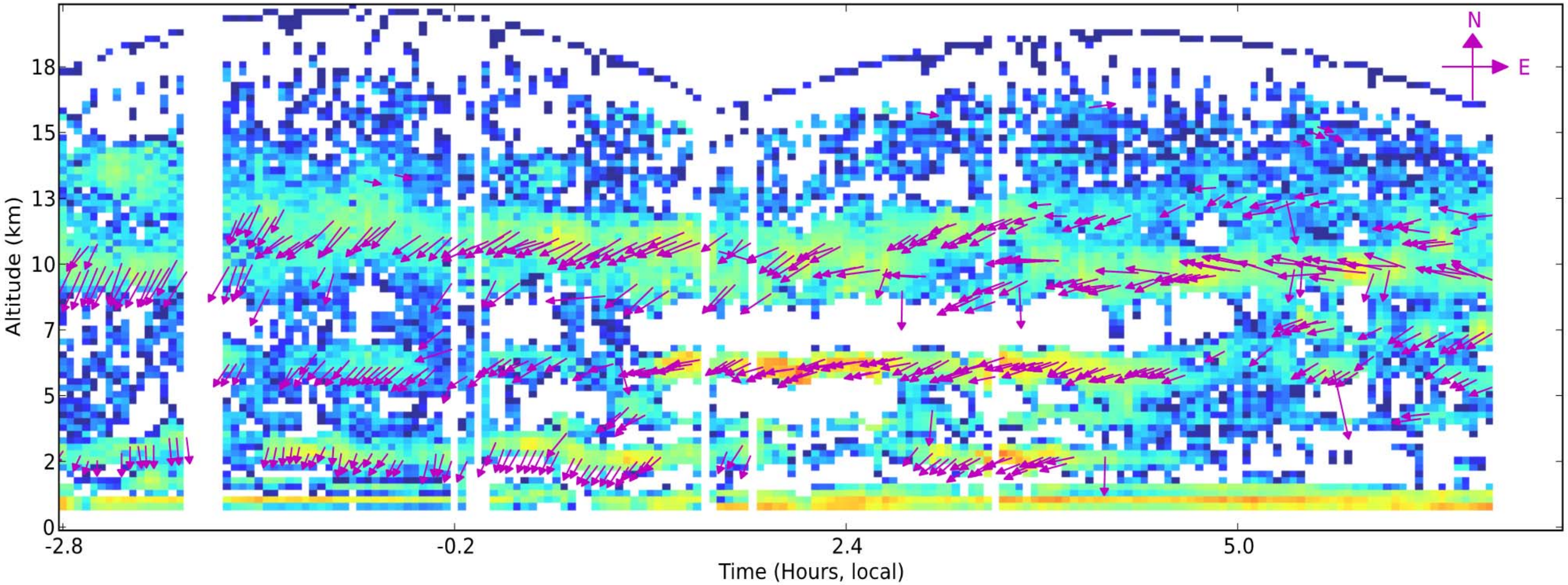
Optical turbulence changes altitude and strength

Osborn et al. AO4ELT3, 2013
Shepherd et al. MNRAS 437(4), 3568-3577, 2013

Results from Stereo-SCIDAR

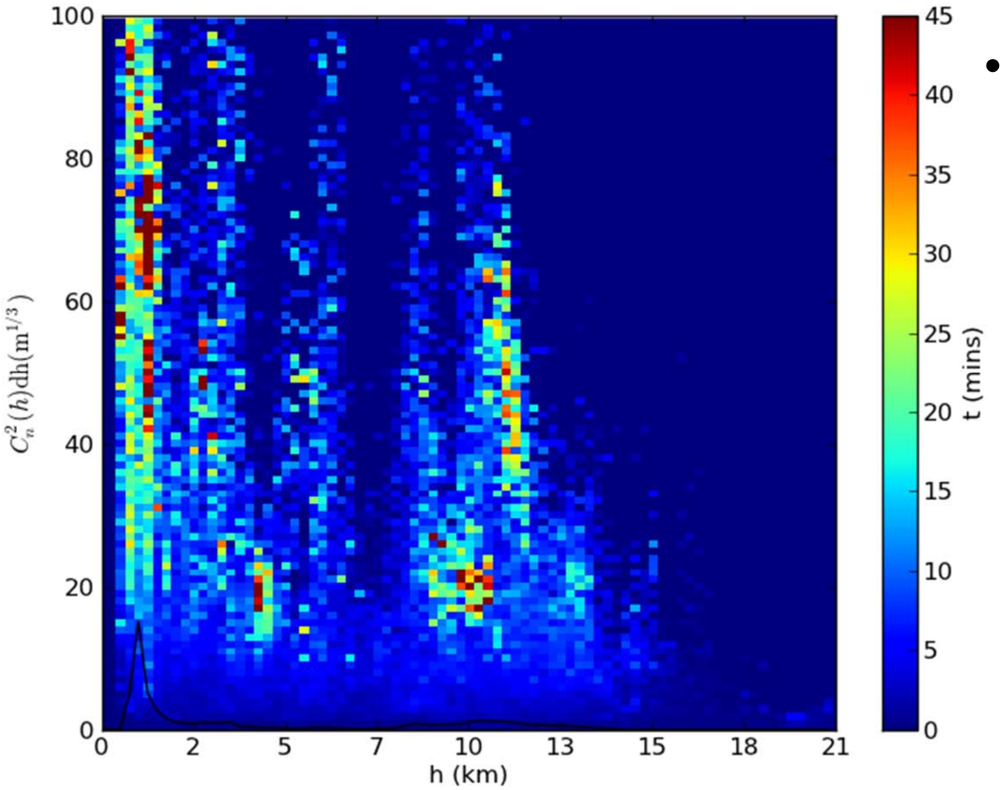
13/09/2013, JKT, La Palma

Subset of wind vectors shown



Optical turbulence velocity, direction and speed, also change

Variability

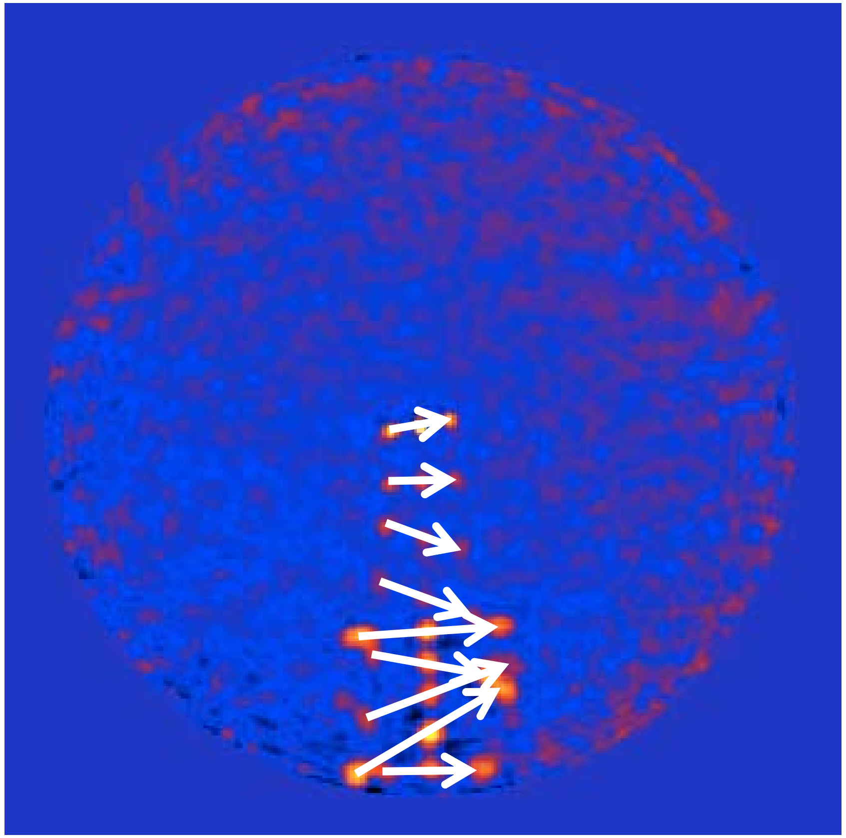


- Average time for a turbulent layer to change by 50 %

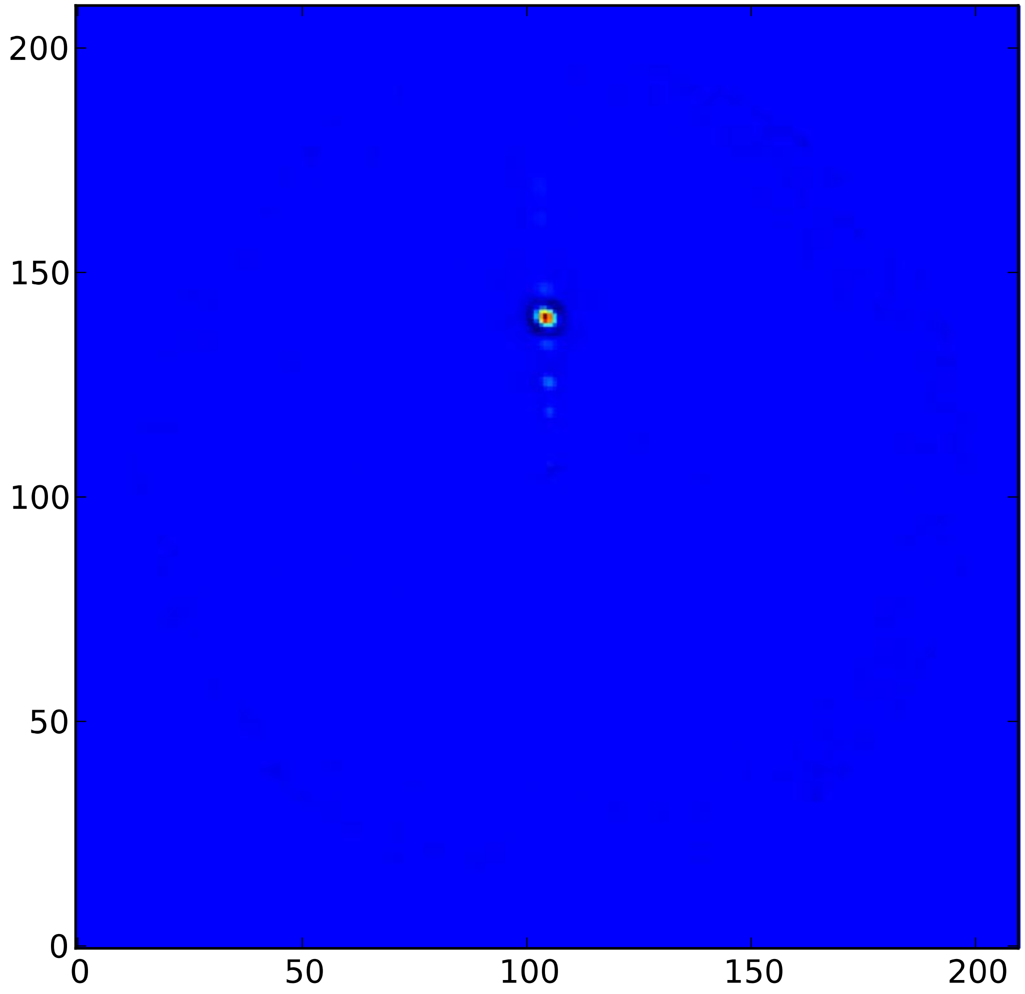
Wind Velocity Profile

Wind profile 25/09/12, NOT 2.5 m, 9 Layers
Cross correlations of sum of three time steps

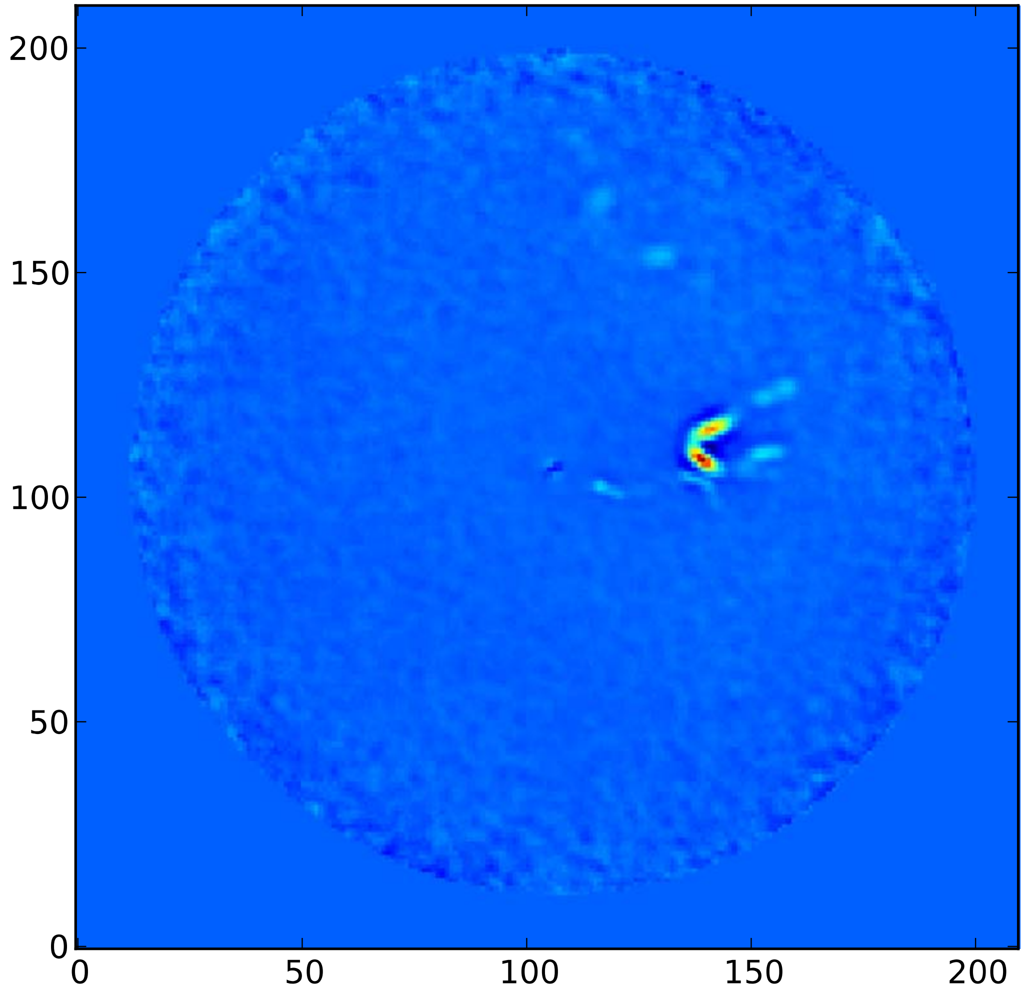
- Cross correlations only contain one set of peaks.
- Peak finding using modified Laplacian of Gaussian algorithm
- Geometric wind profile algorithm rather than wavelet analysis
  - both assume frozen flow
- Use velocity dispersion to separate layers at sub-Fresnel radius level
  - High-vertical resolution SCIDAR
  - Altitude resolution now limited (?) to size of pixels



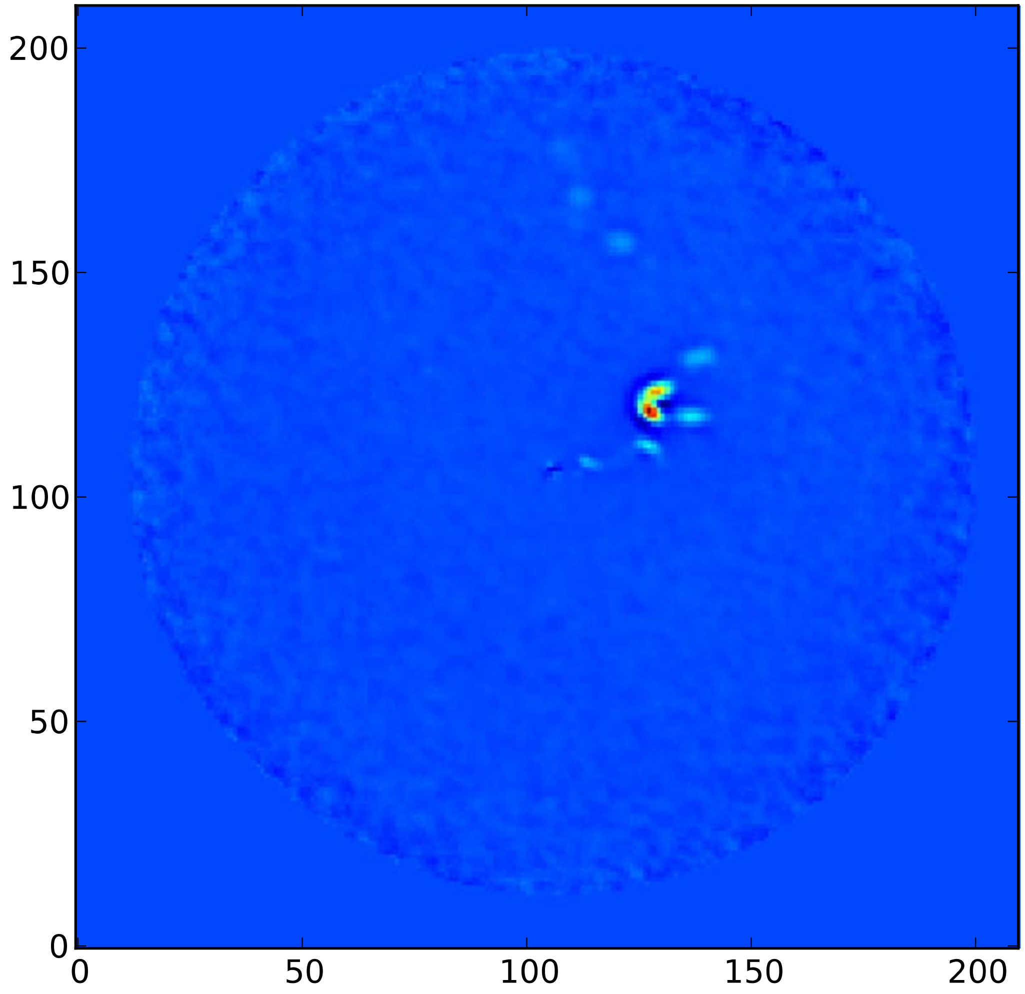
High-vertical resolution SCIDAR



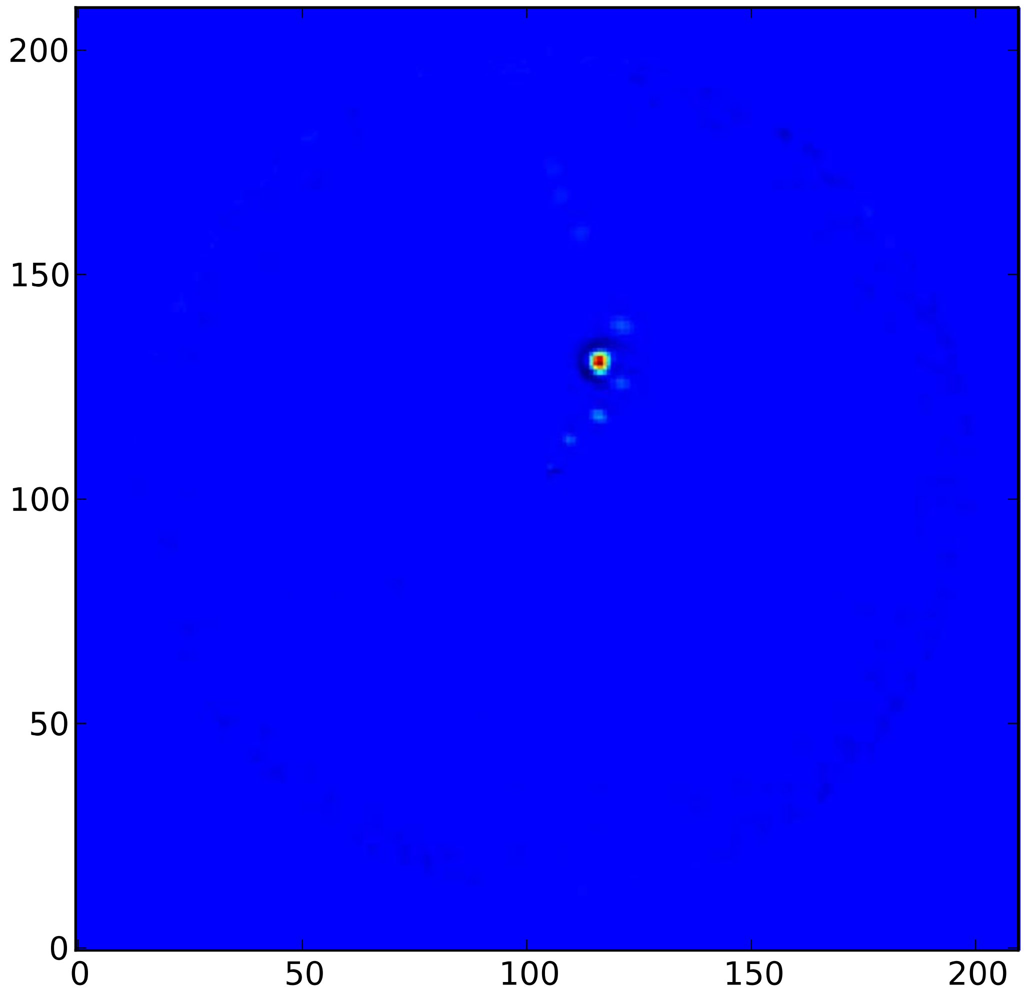
High-vertical resolution SCIDAR



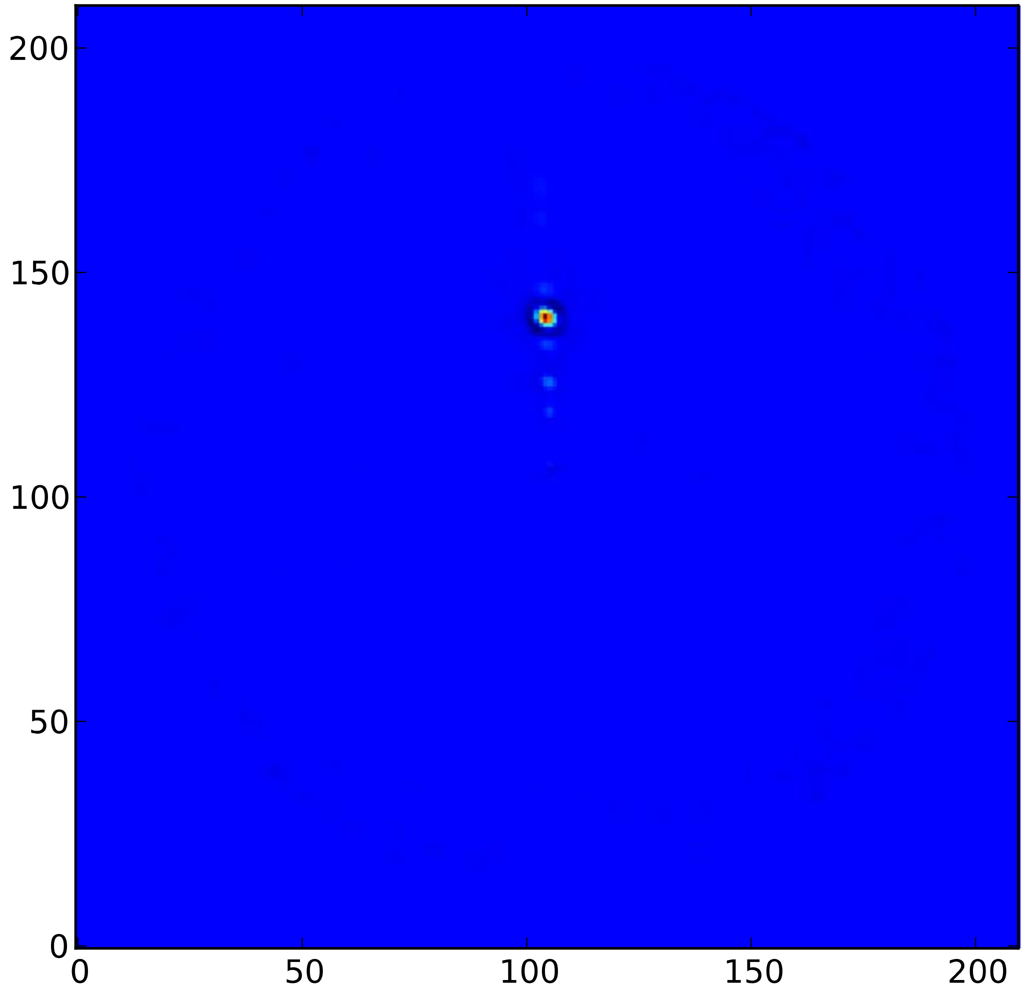
High-vertical resolution SCIDAR



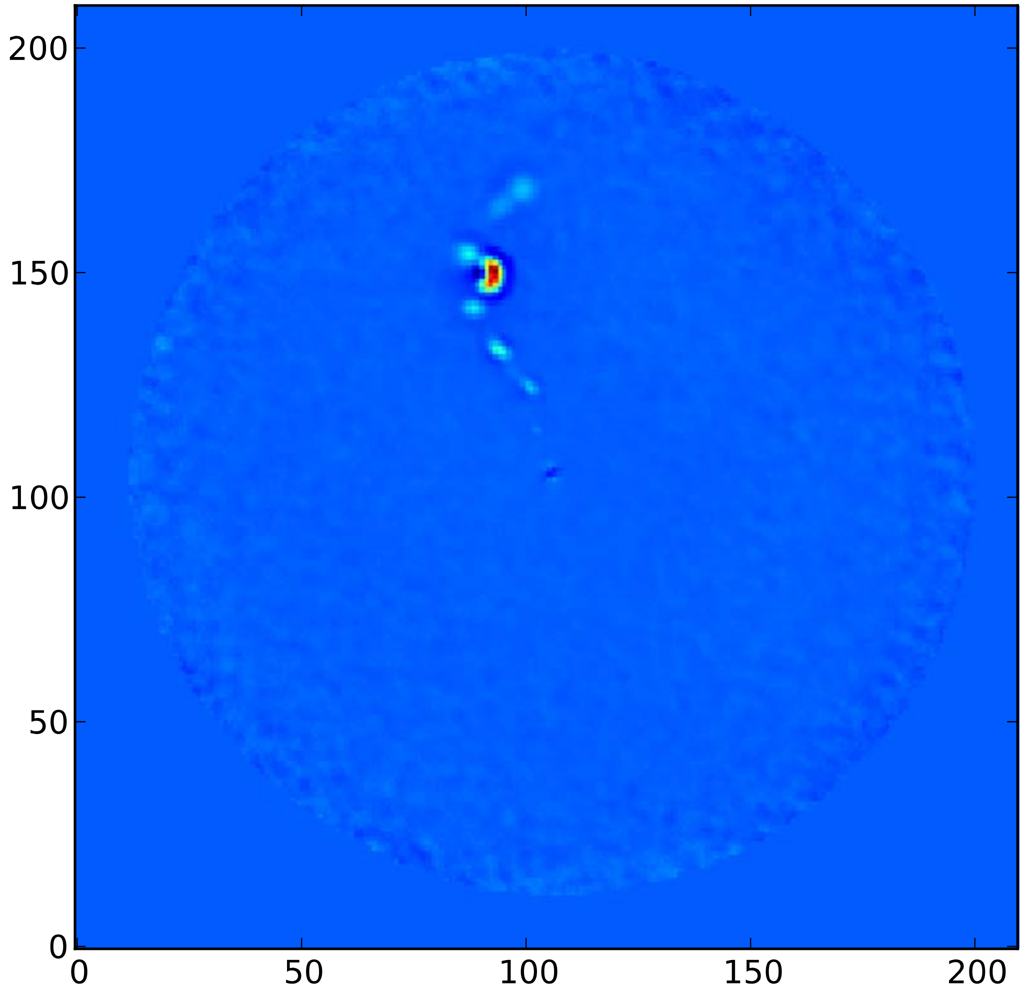
High-vertical resolution SCIDAR



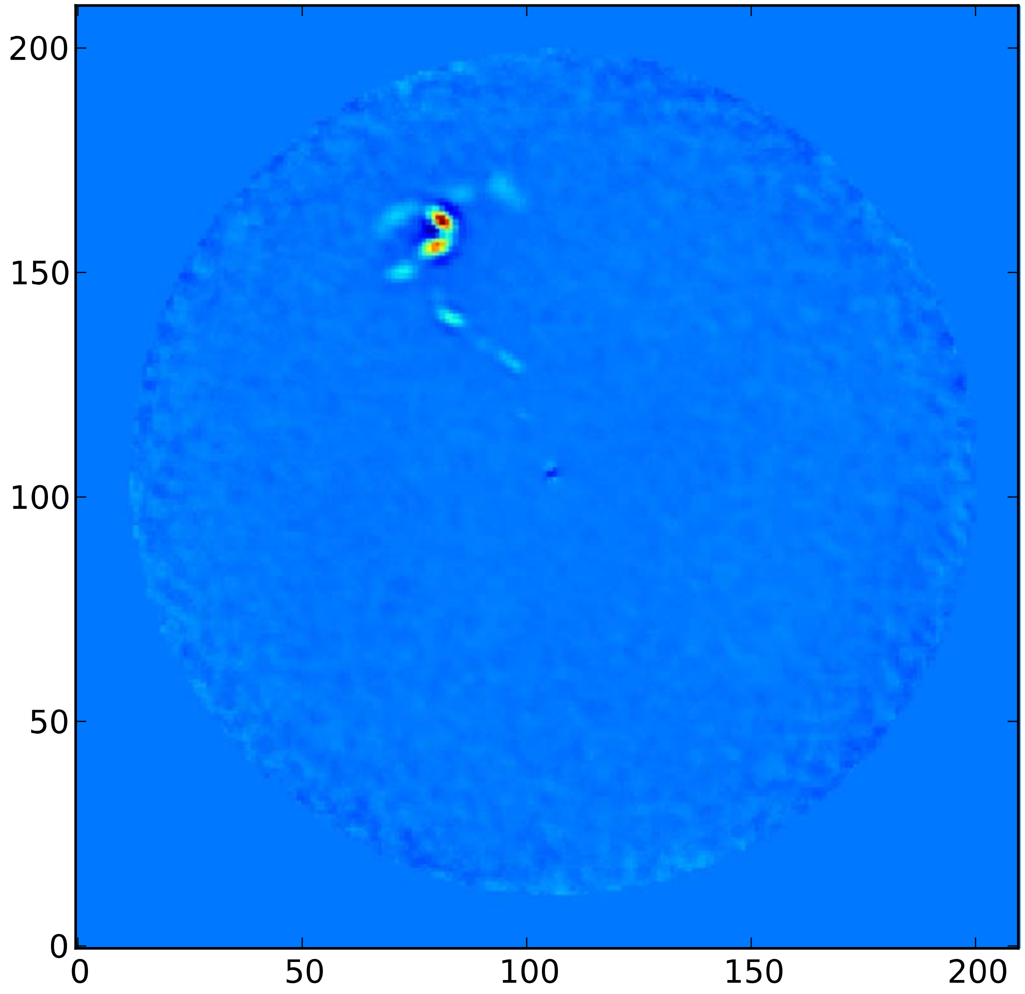
High-vertical resolution SCIDAR



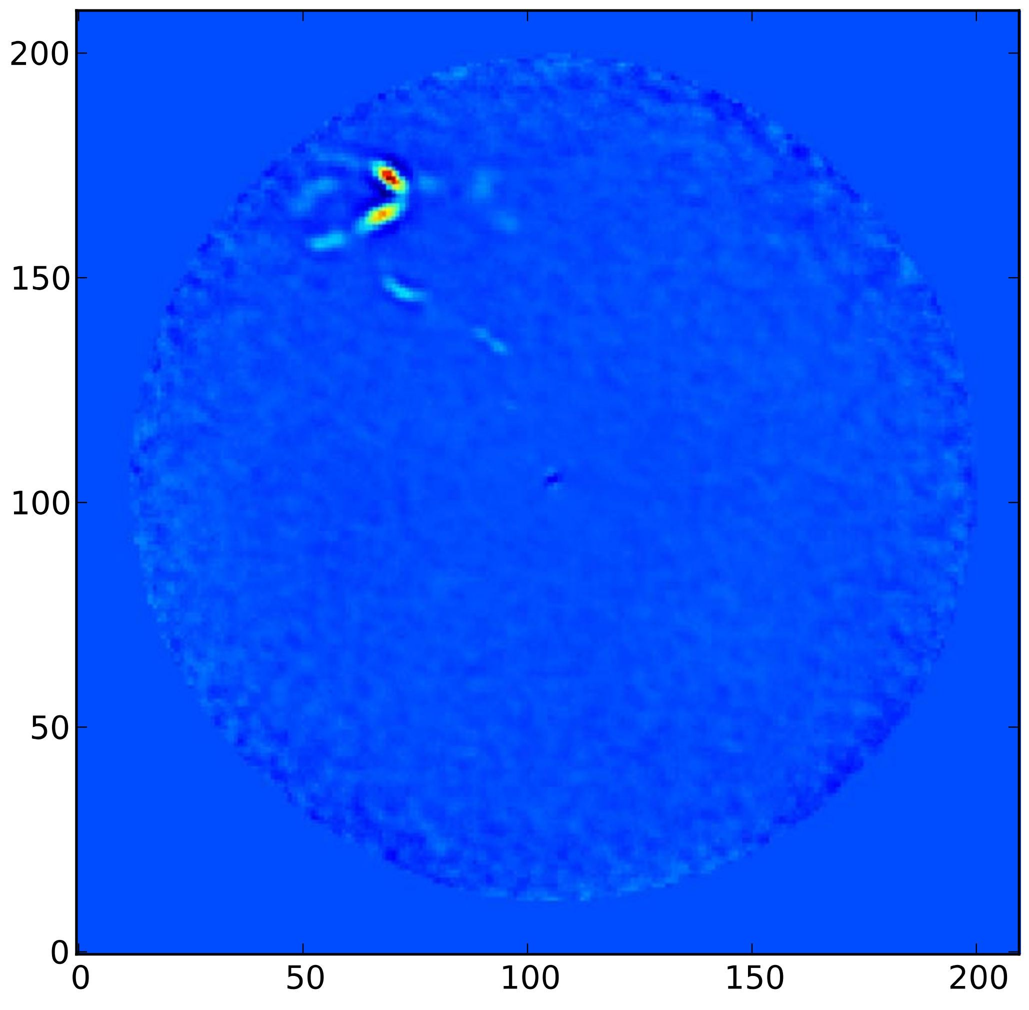
High-vertical resolution SCIDAR



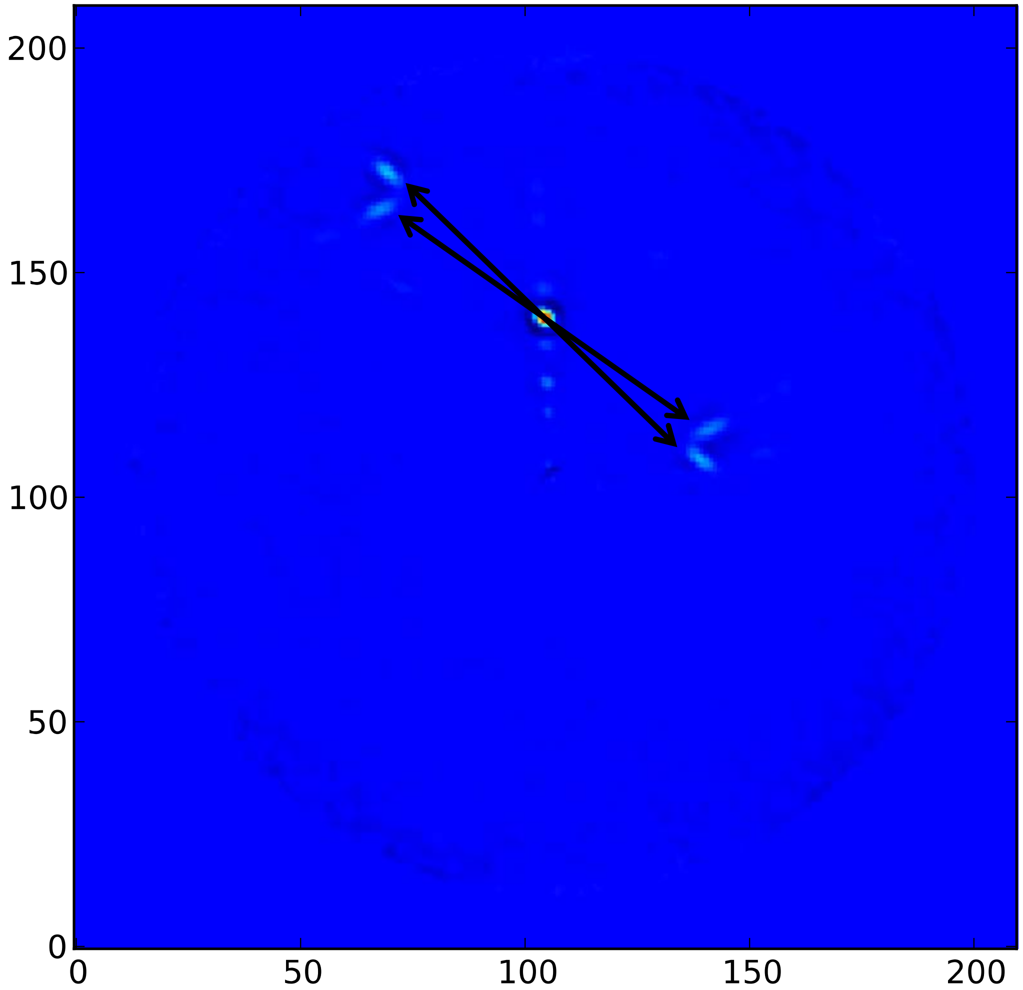
High-vertical resolution SCIDAR



High-vertical resolution SCIDAR



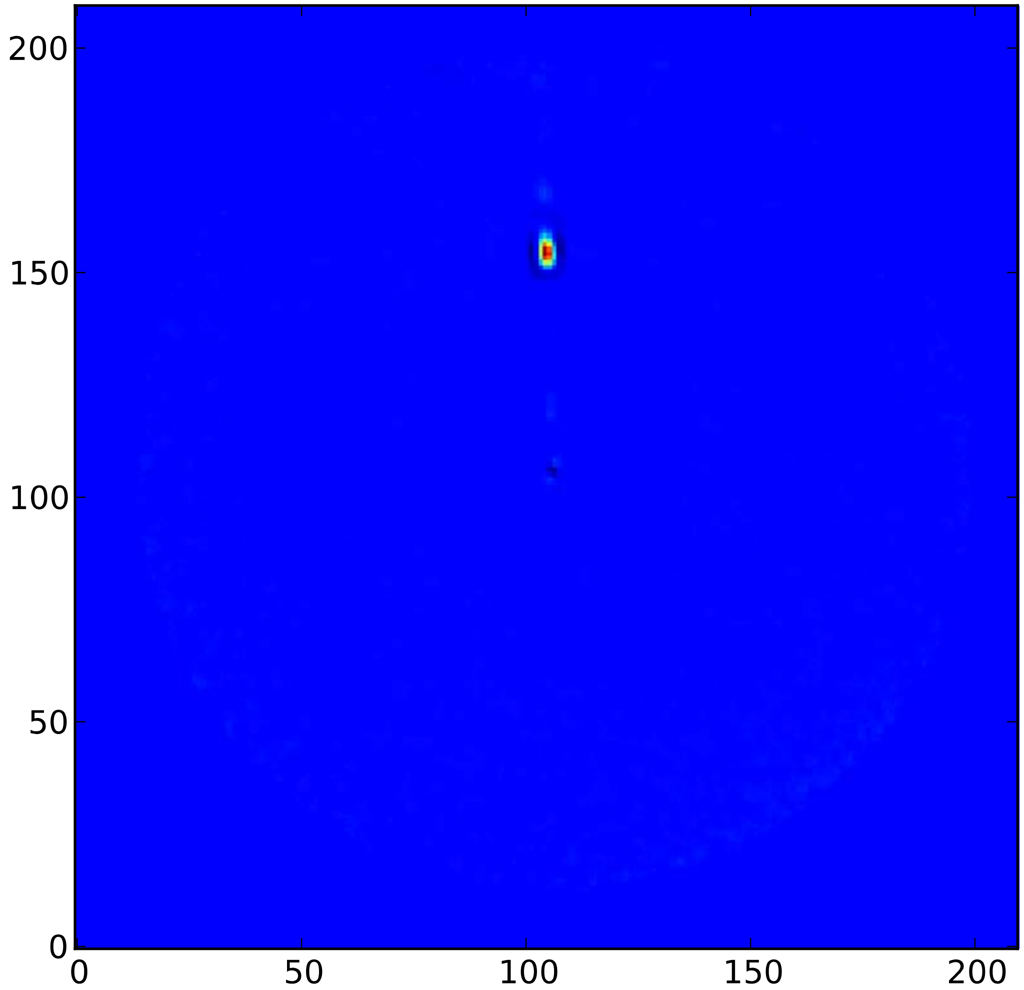
High-vertical resolution SCIDAR



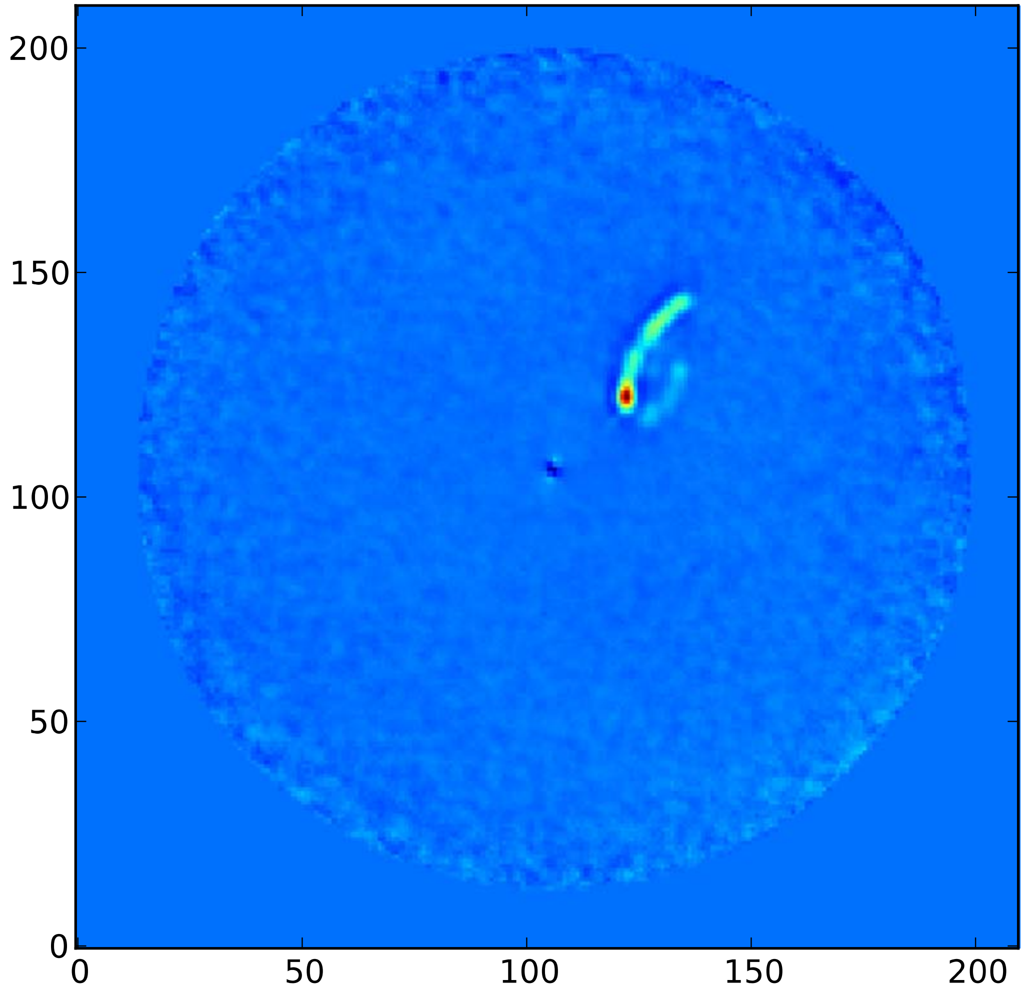
INT, 2.5 m
2014/03/16
2 layers at same altitude?:
h = 8800 m

layer 1	17 m/s	270 deg
layer 2	19 m/s	265 deg

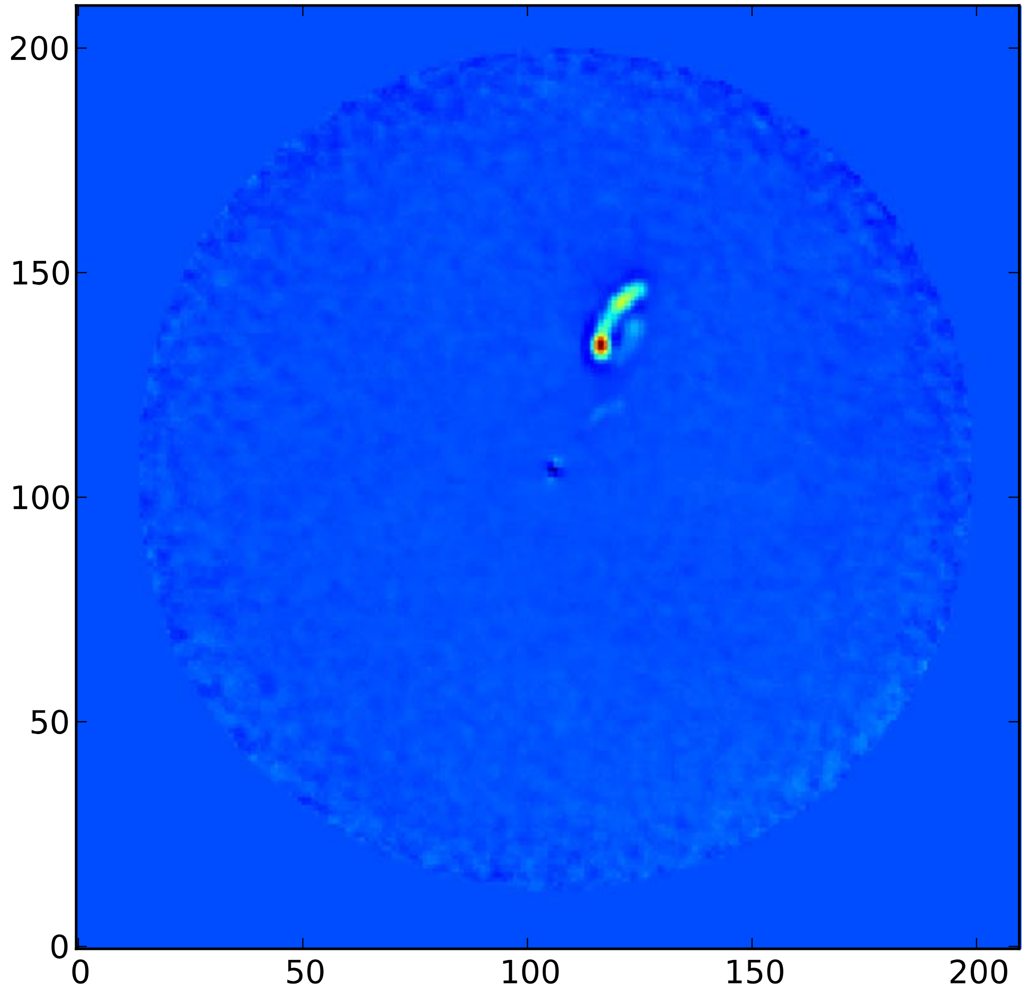
Velocity Dispersion



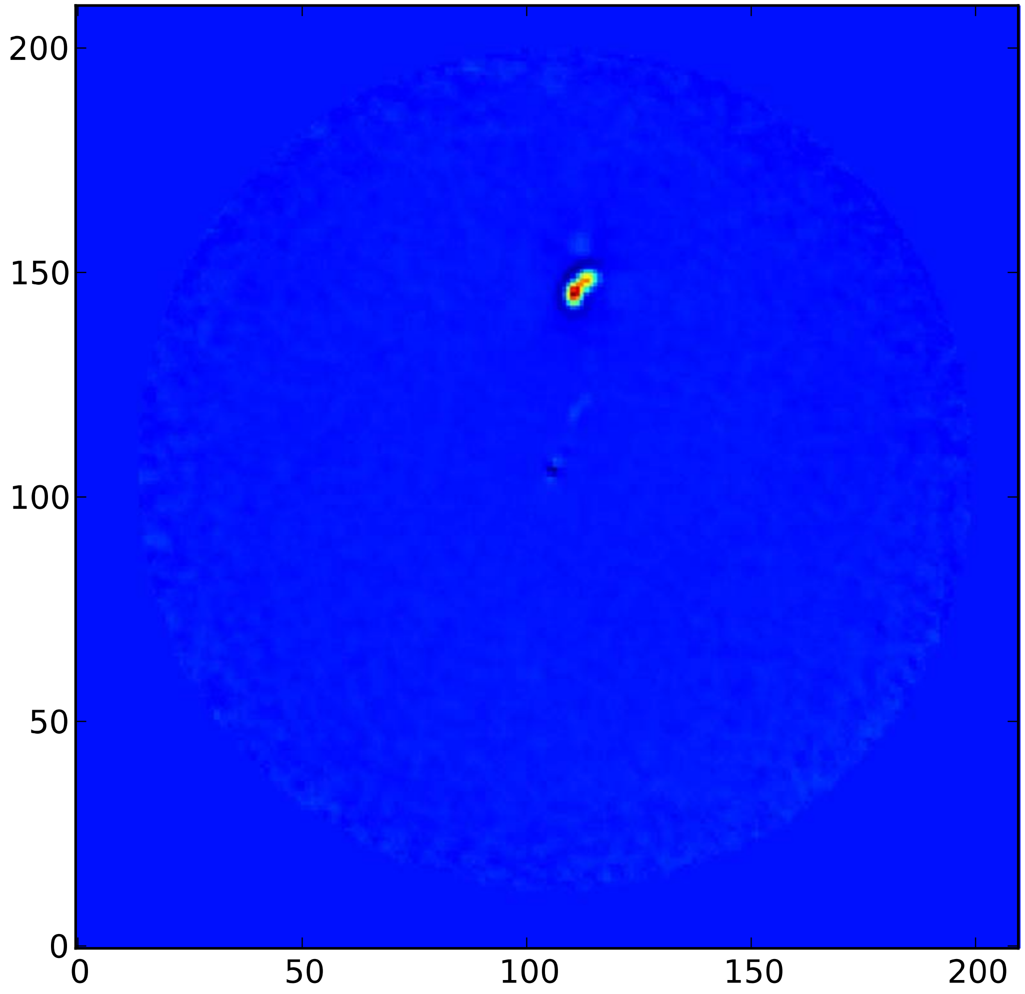
Velocity Dispersion



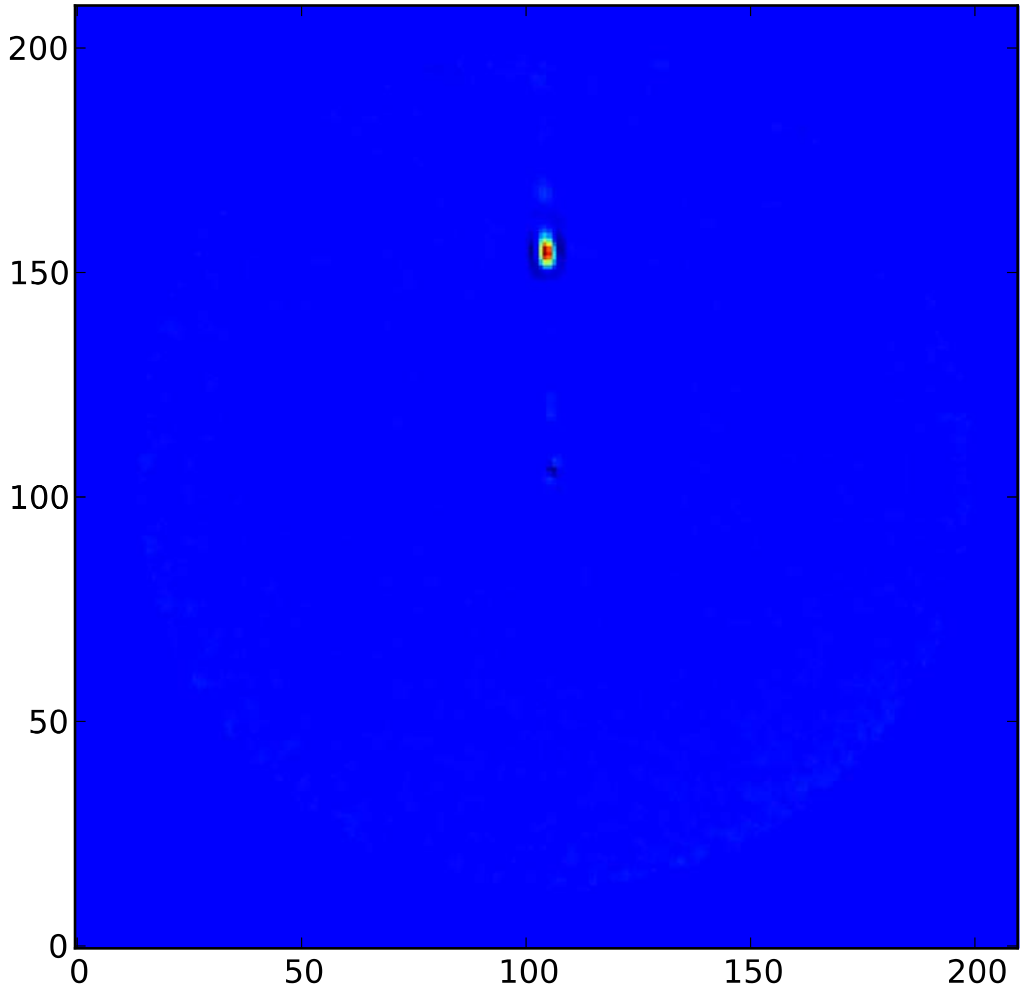
Velocity Dispersion



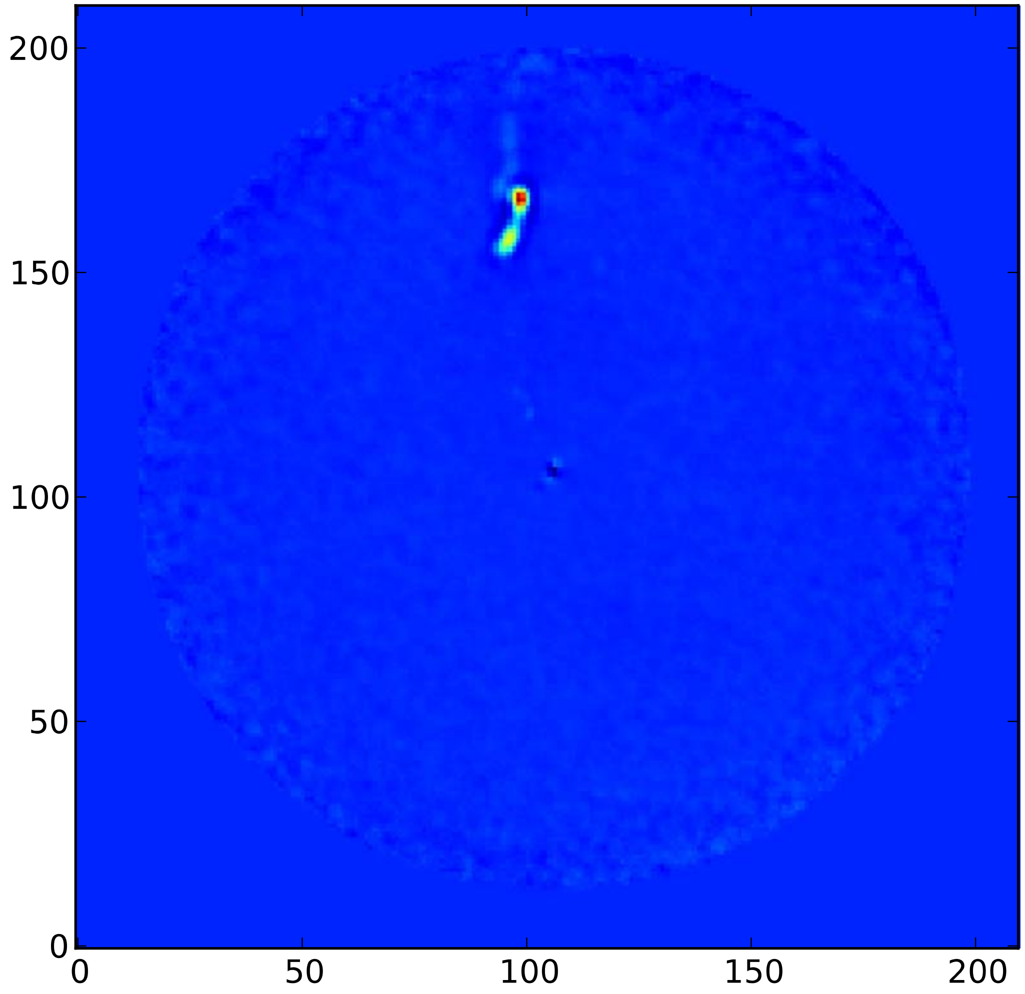
Velocity Dispersion



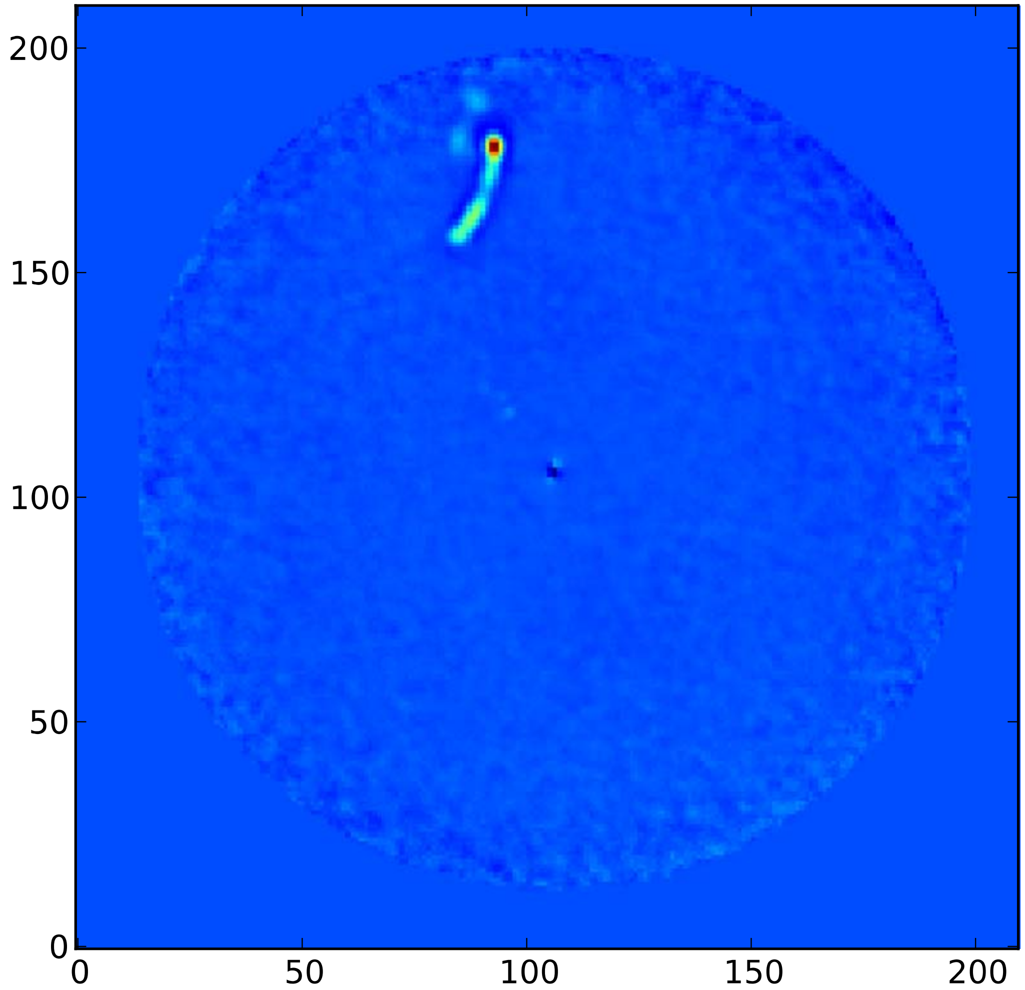
Velocity Dispersion



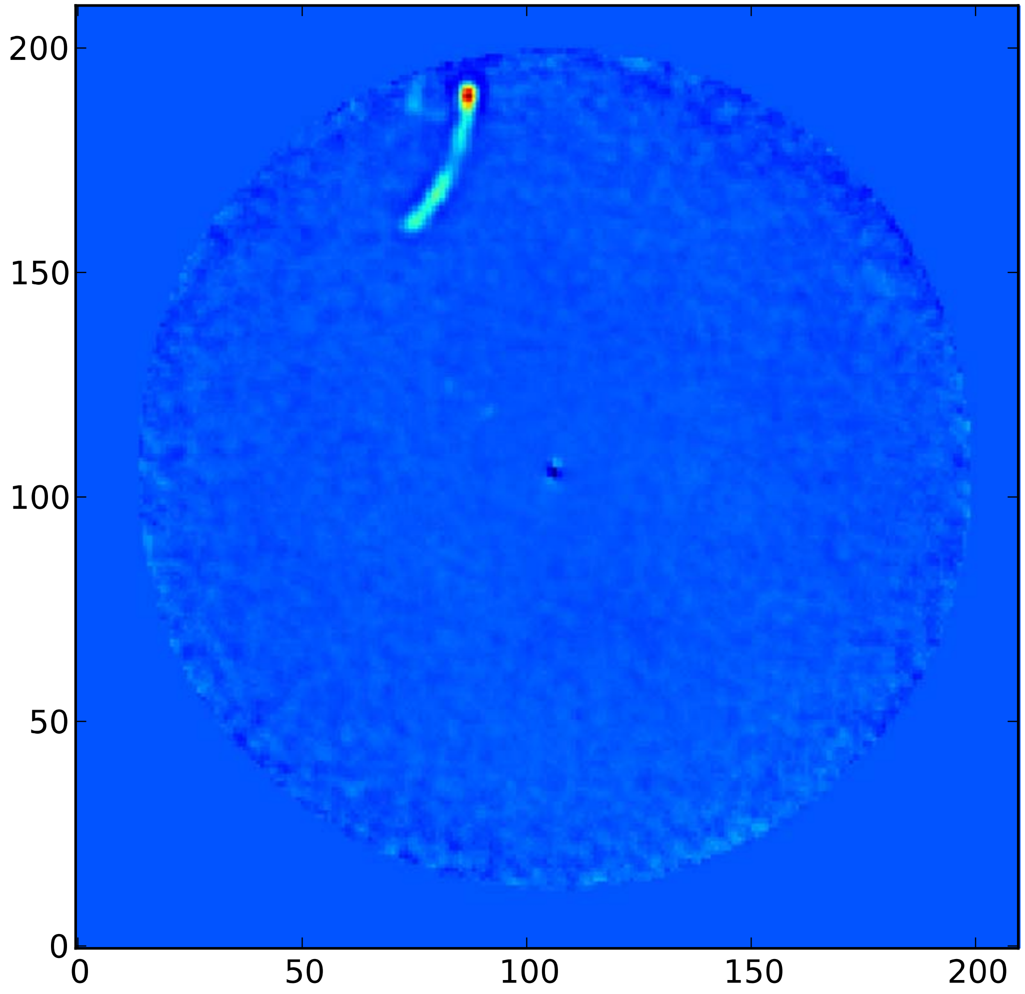
Velocity Dispersion



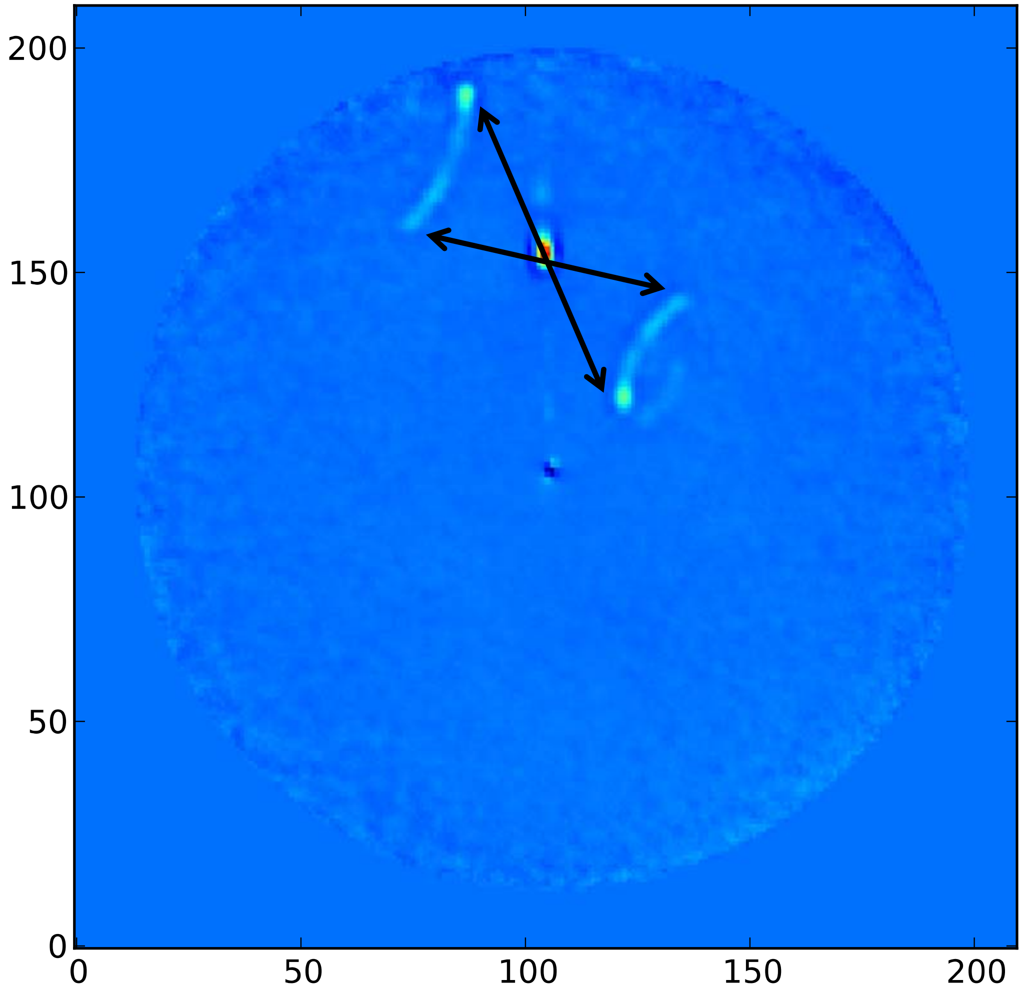
Velocity Dispersion



Velocity Dispersion



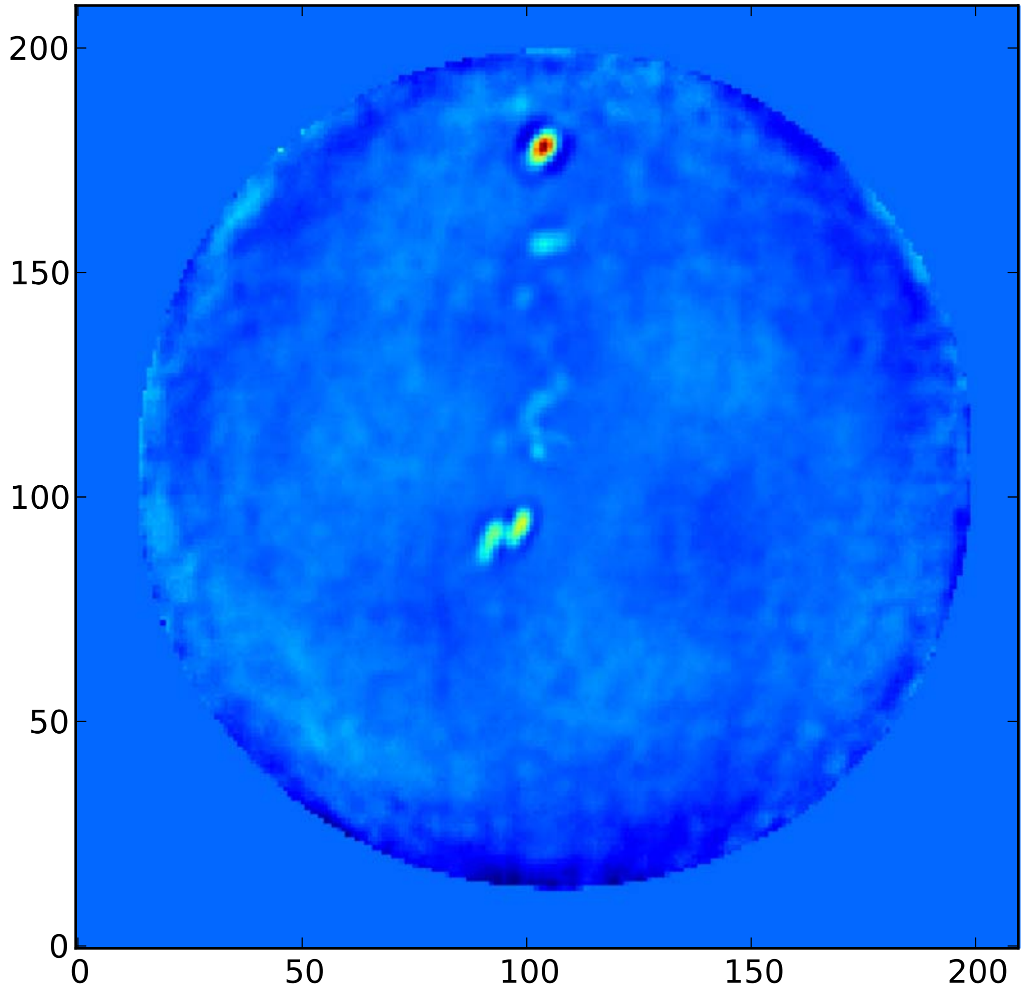
Velocity Dispersion



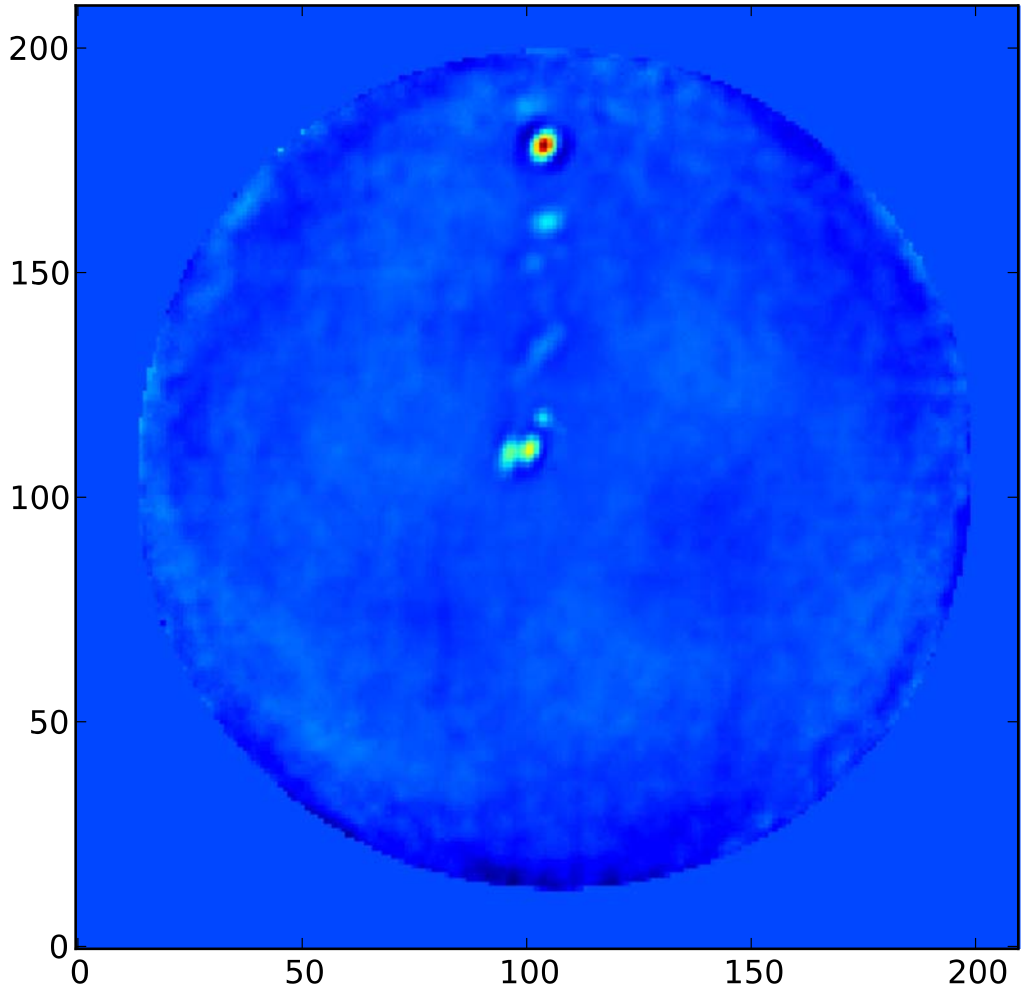
INT, 2.5 m
2014/03/17

- Velocity dispersion within turbulent layer?
- Strong turbulence and low noise
  - see structures of turbulent layers

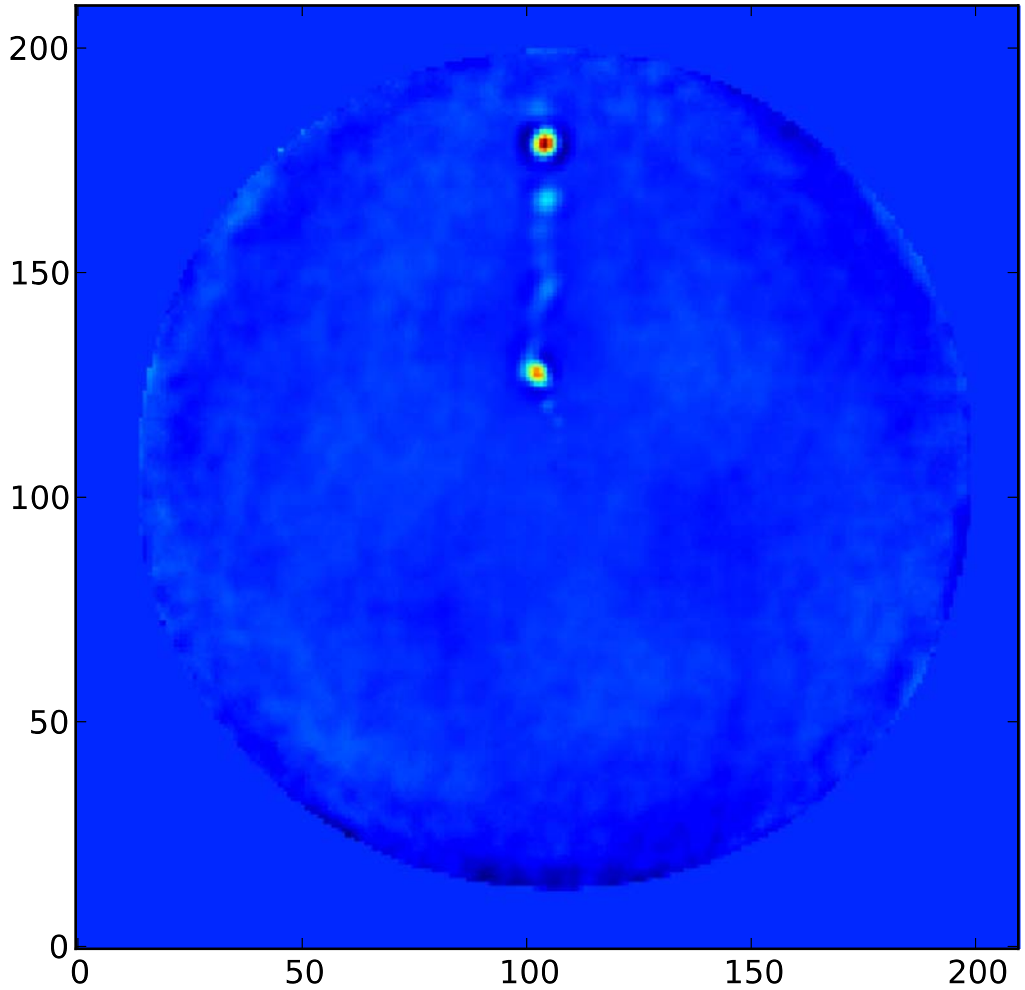
Slow high altitude turbulence



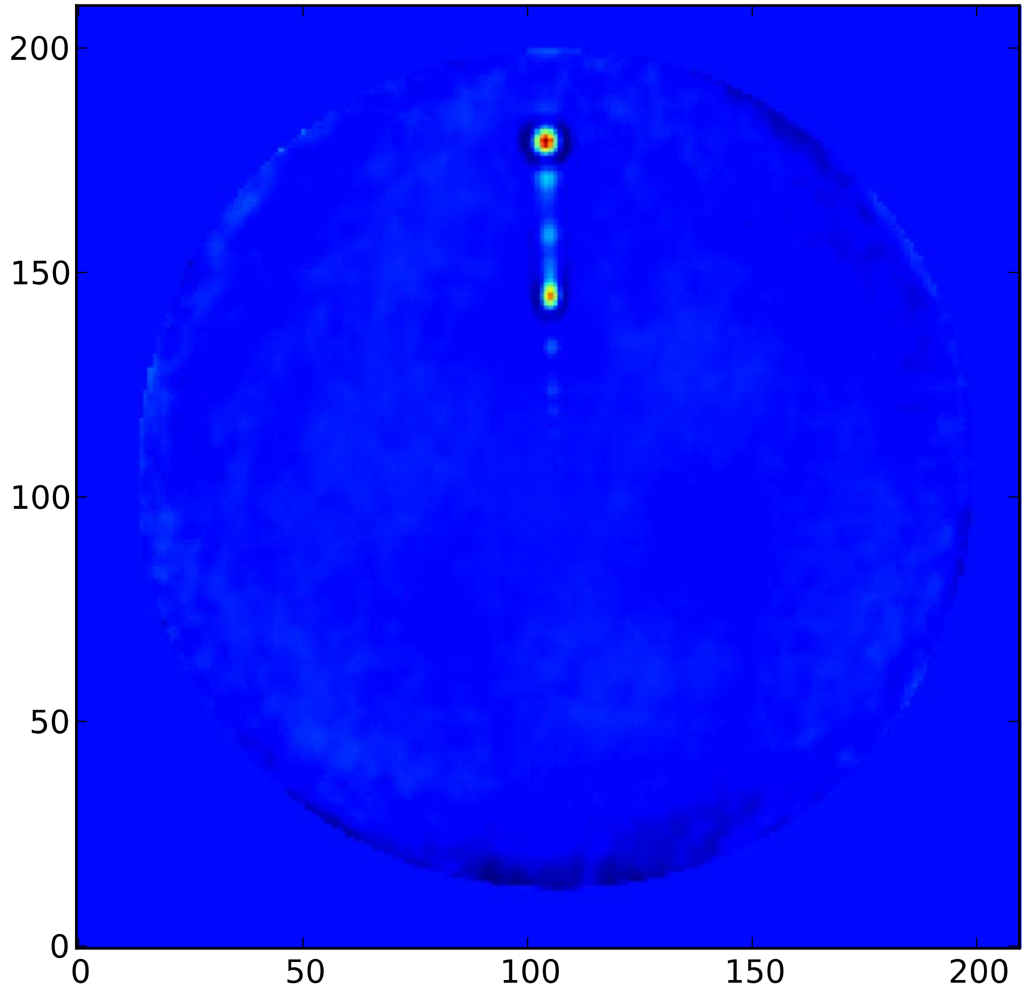
Slow high altitude turbulence



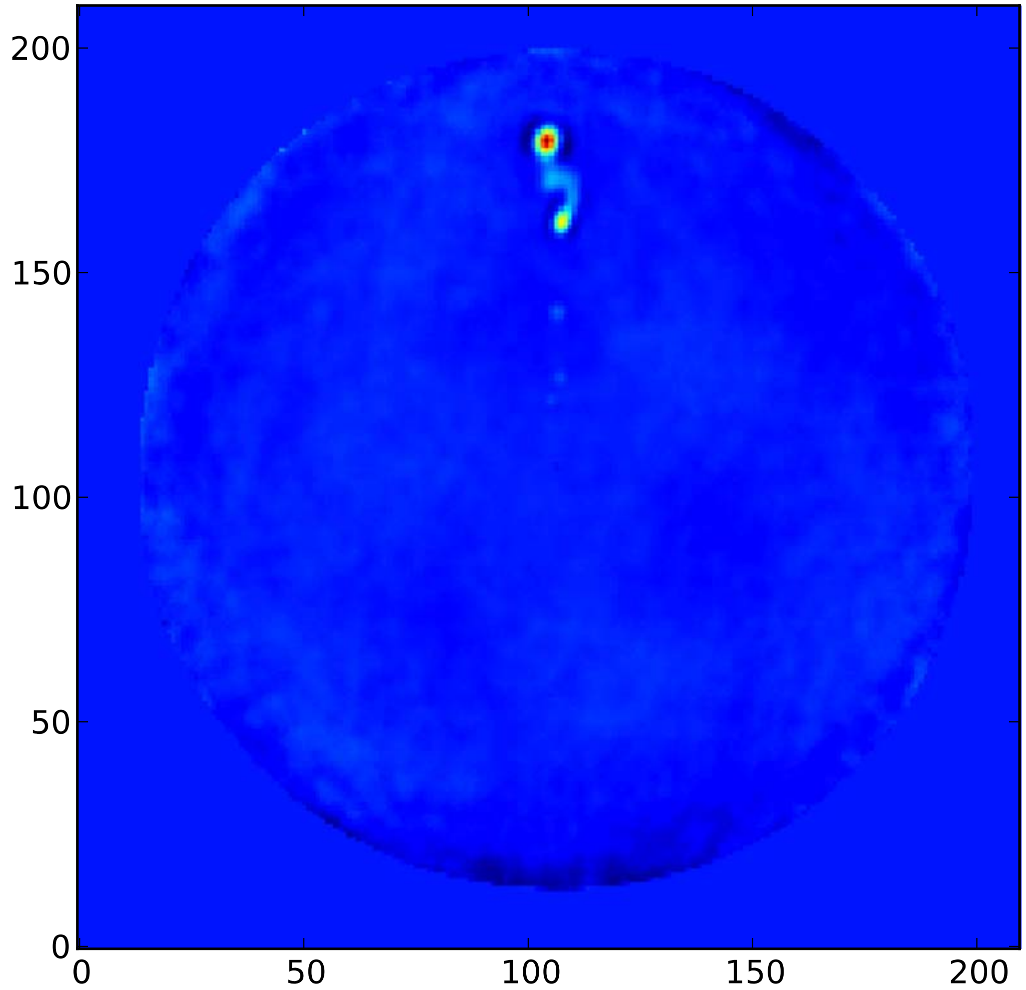
Slow high altitude turbulence



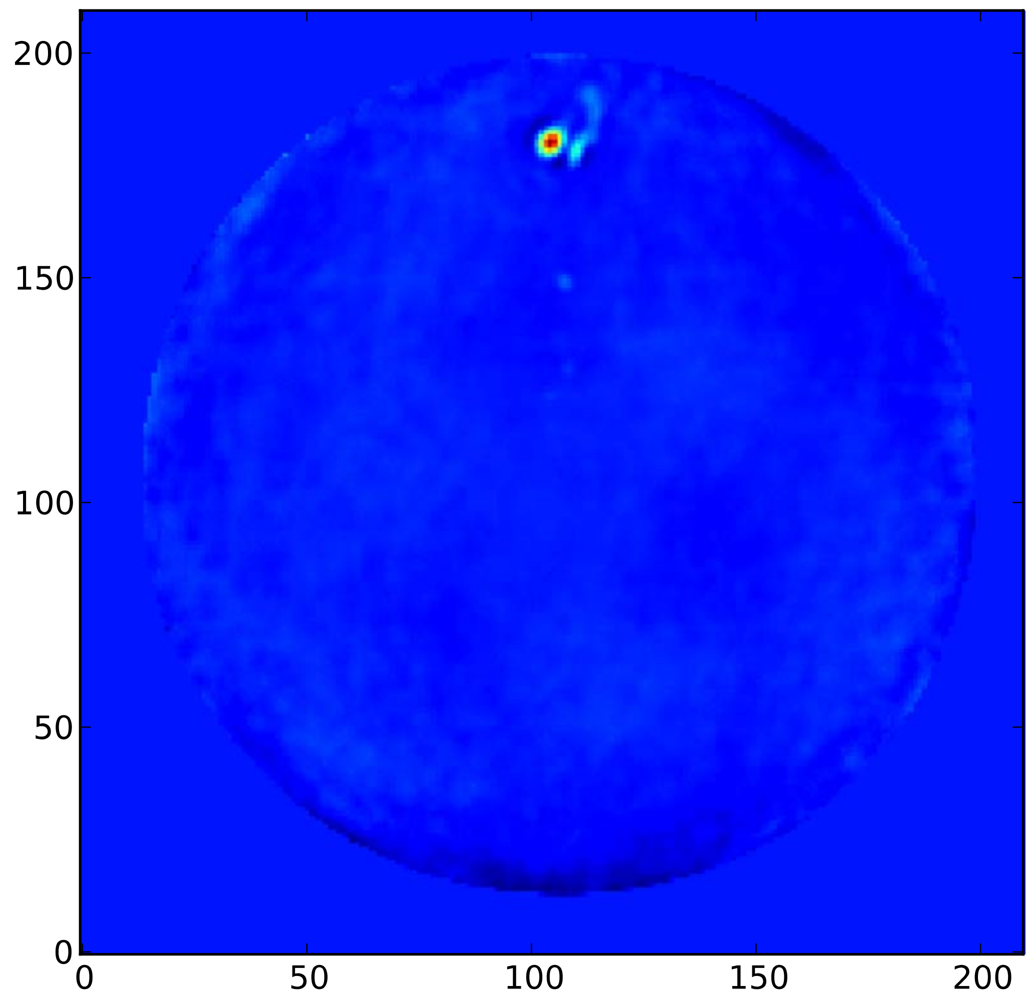
Slow high altitude turbulence



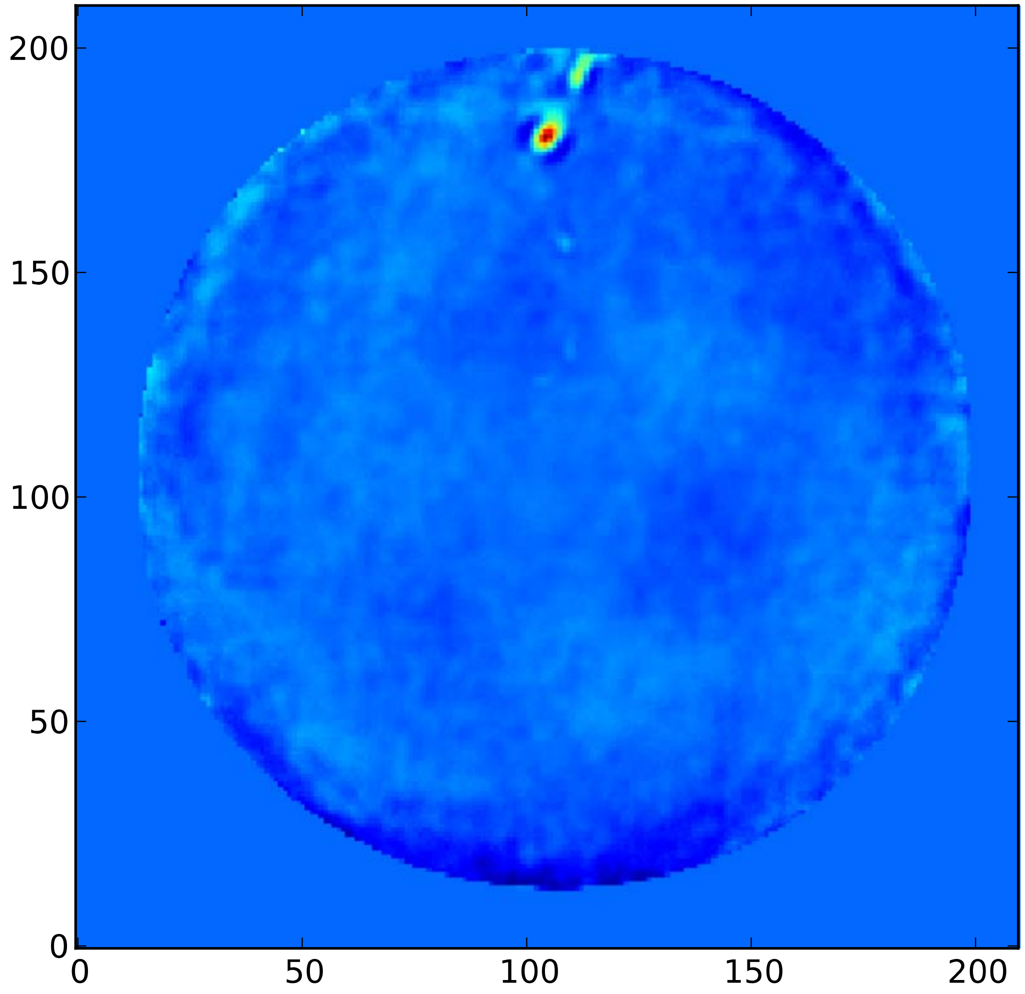
Slow high altitude turbulence



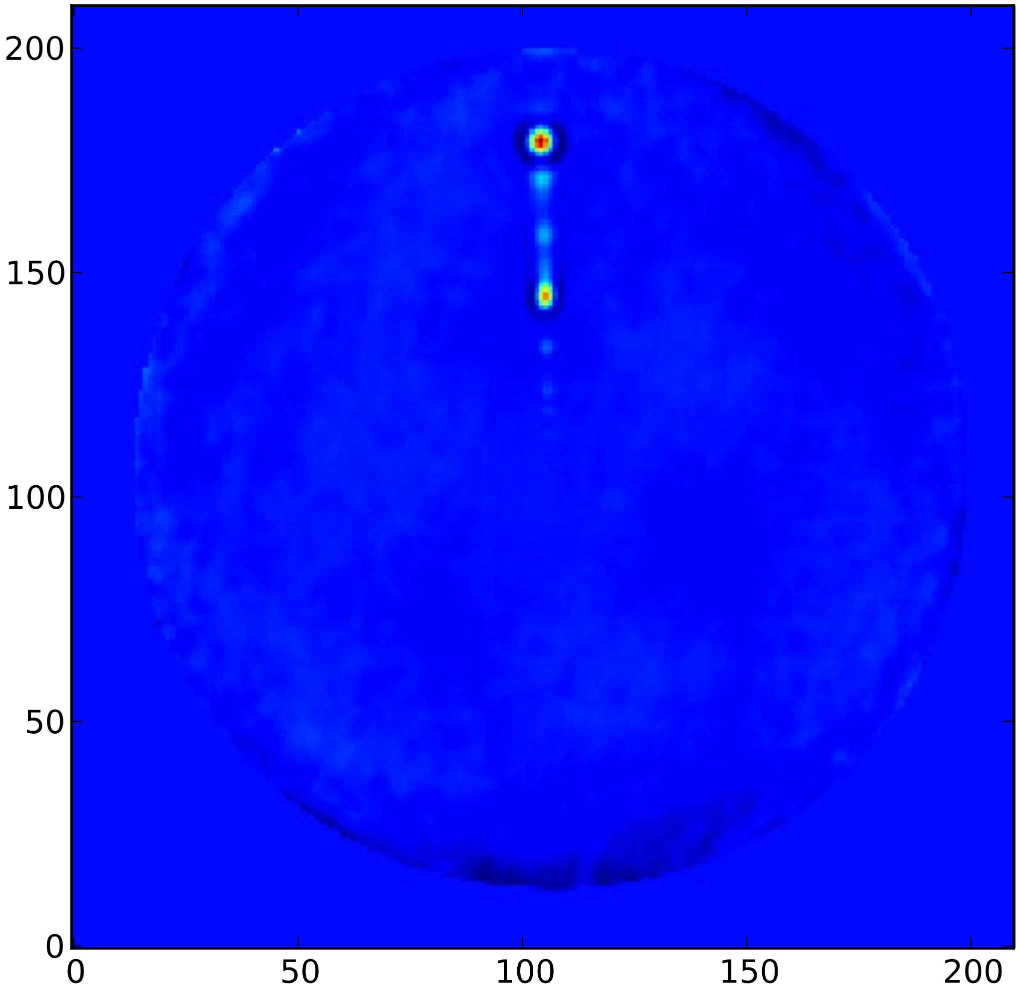
Slow high altitude turbulence



Slow high altitude turbulence



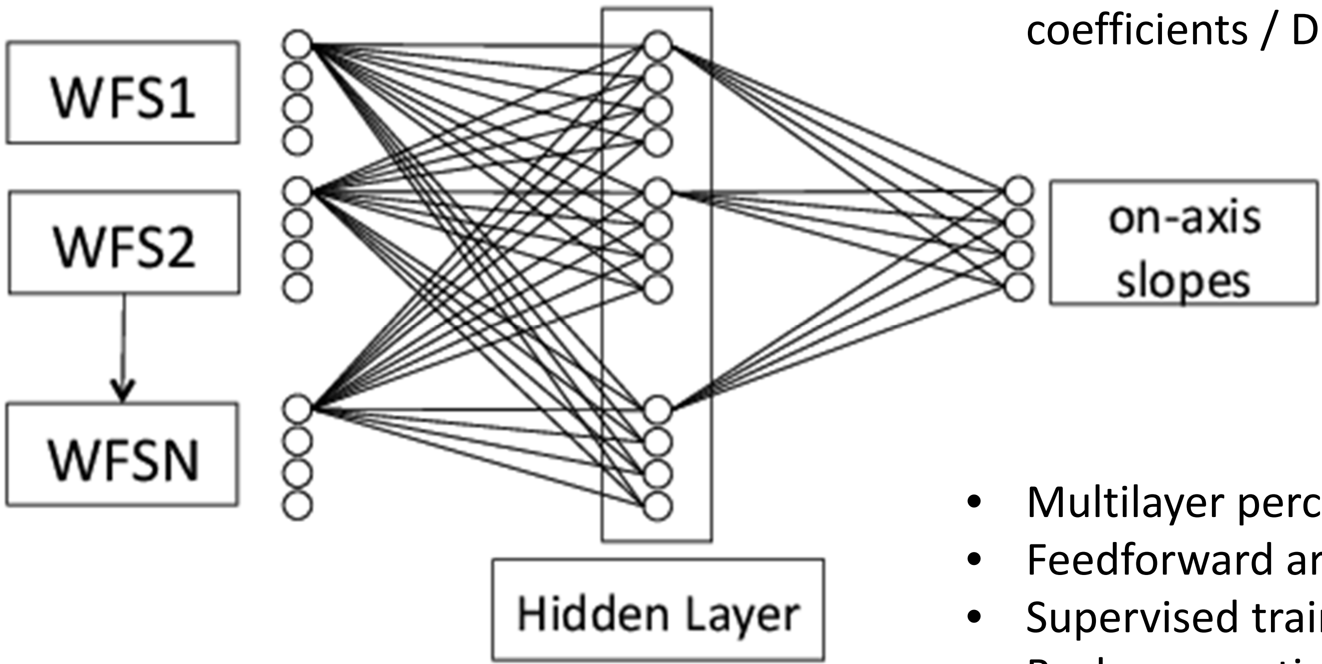
Slow high altitude turbulence



INT, 2.5 m
2014/03/15

- Very slow (<<1m/s) high turbulent layer
- 1 pixel/frame = 2 m/s
- Use to study Taylor frozen Flow

Artificial Neural Networks for Open-loop tomography



- Trained with off-axis guide star slopes and desired target slopes / Z coefficients / DM commands

- Multilayer perceptron (MLP):
- Feedforward architecture
- Supervised training process
- Backpropagation training algorithm

Training

- Can not train on sky
- Train on optical calibration bench or simulation
- Trained with same asterism that will be used on-sky
- Turbulent phase screen moving in altitude steps of 1/10th subaperture

$$\mathrm{d}h = \frac{1}{10}\frac{D}{\theta n}$$

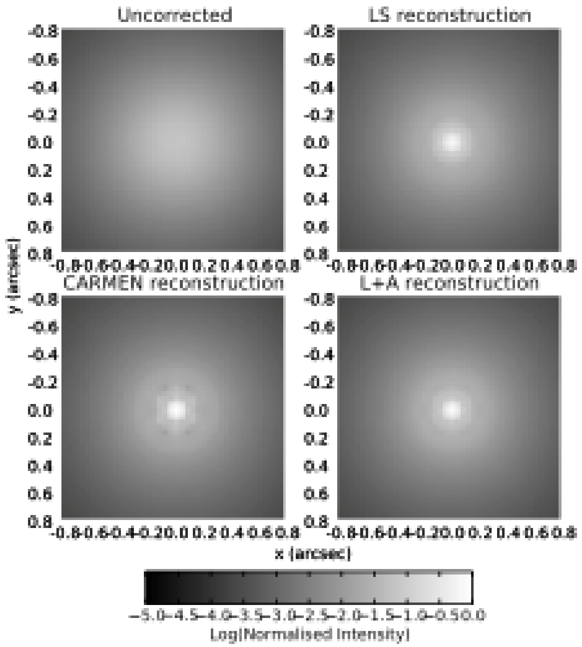
- ANN is exposed to turbulent layers at all altitudes
- Train with off-axis slopes (input) and measured on-axis slopes (desired output)

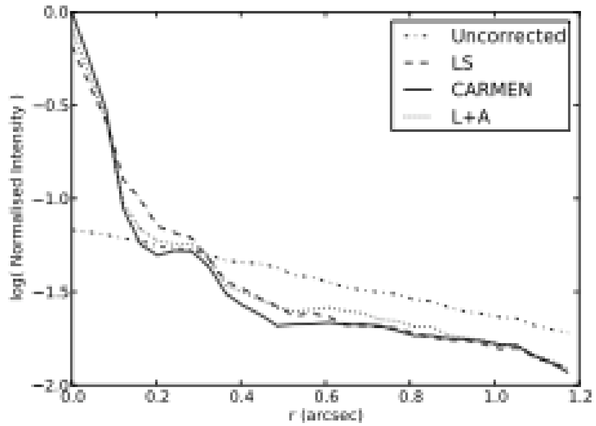
Simulation Results

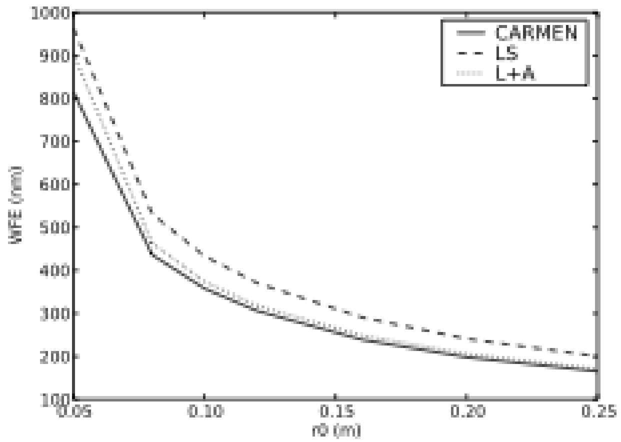
- Trained in simulation
  - with single layer
- Tested with three 4 layer profiles (Good, Median, Bad)

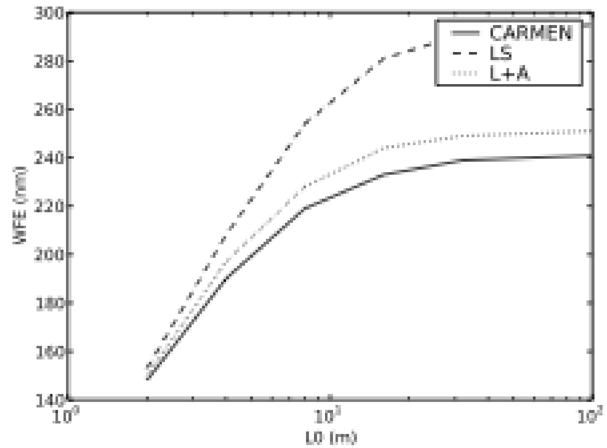
Parameter	Values			Units
Test Name	atm1	atm2	atm3	
r_0	0.16	0.12	0.085	m
Layer 1				
Altitude	0	0	0	m
Relative strength	0.65	0.45	0.80	
Wind Speed	7.5	7.5	10	m/s
Wind direction	0	0	0	degrees
Layer 2				
Altitude	4000	2500	6500	m
Relative strength	0.15	0.15	0.05	
Wind Speed	12.5	12.5	15	m/s
Wind direction	330	330	330	degrees
Layer 3				
Altitude	10000	4000	10000	m
Relative strength	0.10	0.30	0.10	
Wind Speed	15	15	17.5	m/s
Wind direction	135	135	135	degrees
Layer 4				
Altitude	15500	13500	15500	m
Relative strength	0.10	0.10	0.05	
Wind Speed	20	20	25	m/s
Wind direction	240	240	240	degrees

Simulations

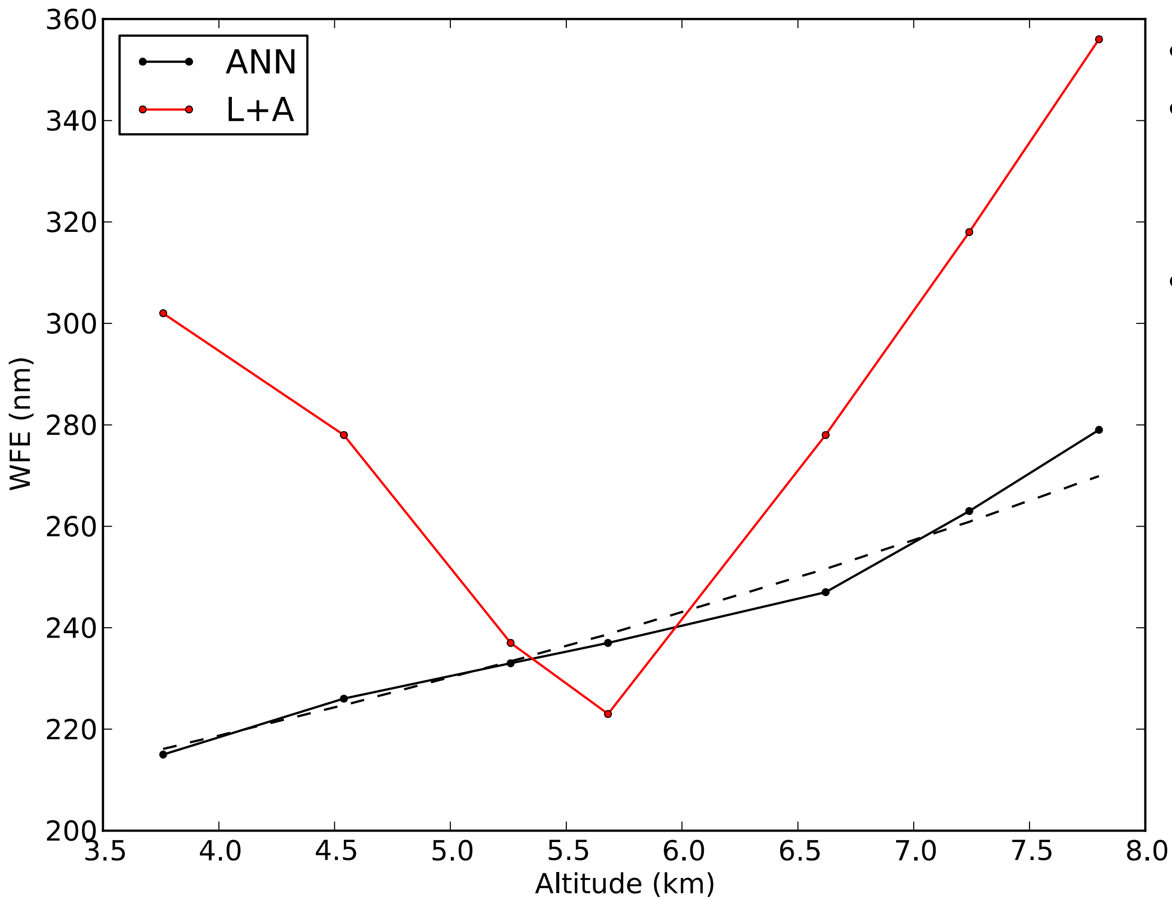






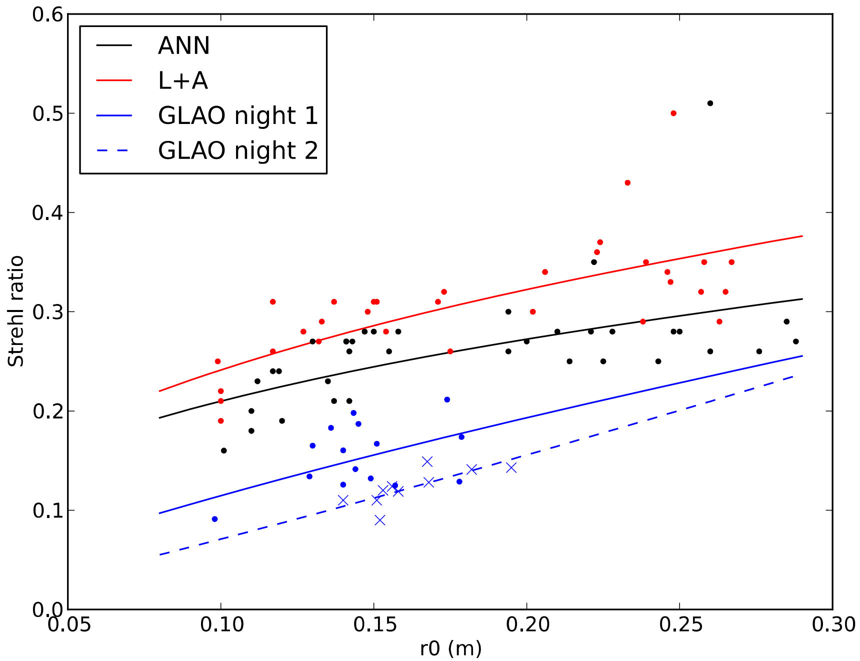


Lab Bench Results

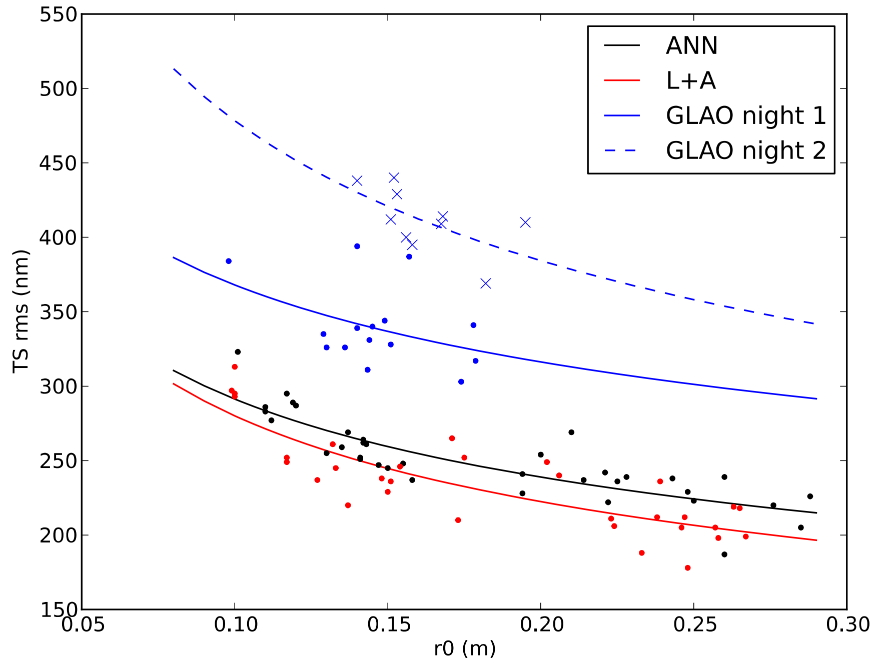


- Trained on bench
- One phase screen at the ground and one moving up in altitude
- ANN error increase due to pupil area overlap decreasing

On-sky Results

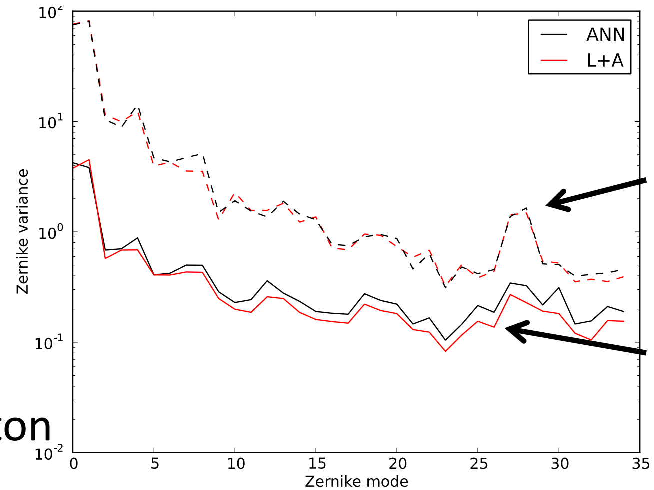


Strehl ratio



Open-loop WFE

- Trained on bench
- Difference at higher orders

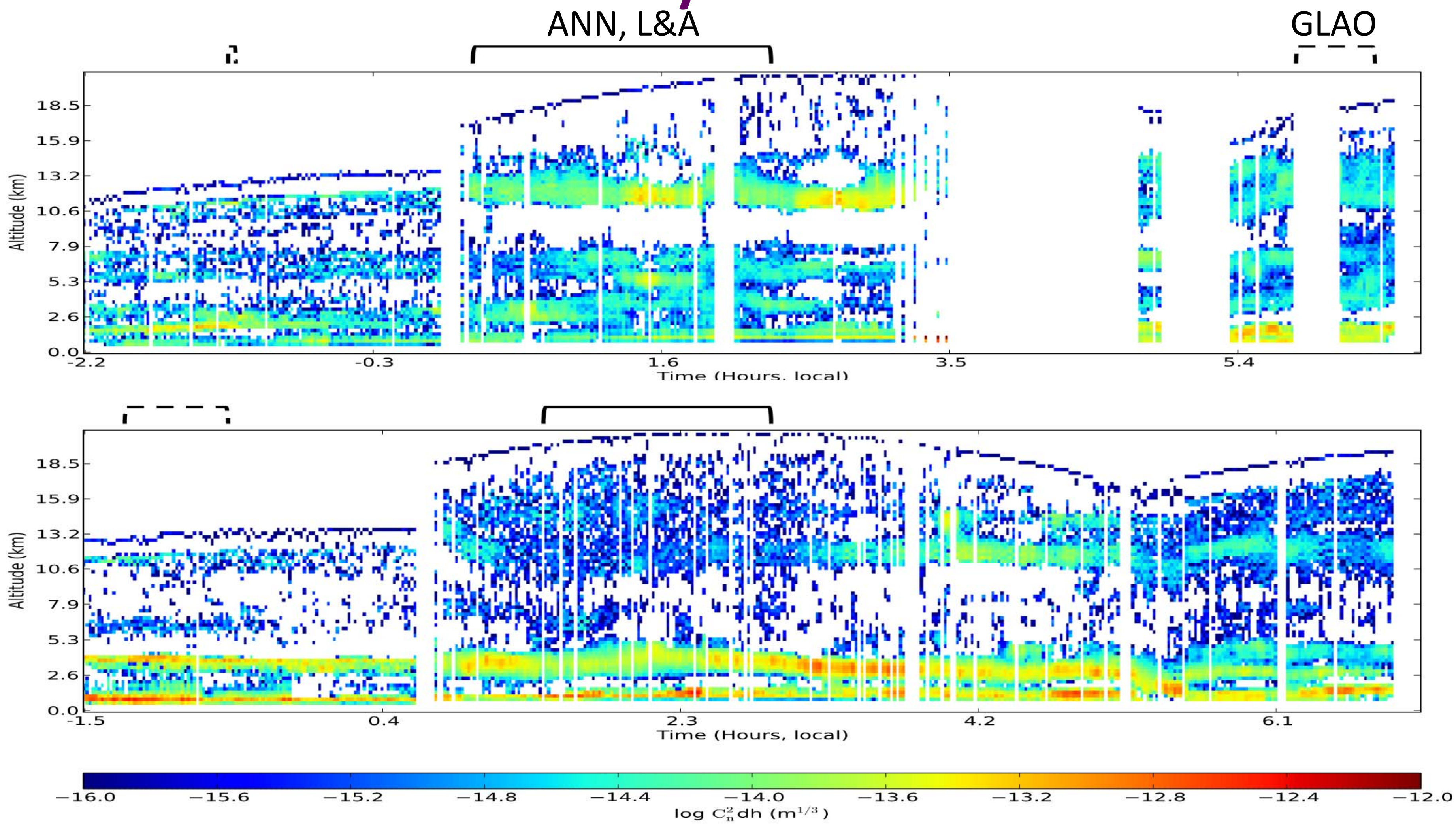


Zernike decompositon

Open-loop (average of off-axis guide stars)

Closed-loop (from corrected truth sensor)

On-sky results



- Turbulence profile different at GLAO times but still turbulence above ground -> Tomographic reconstruction and not just good GLAO

Discussion

- Known issues:
  - Training with bench phase screens that do not converge statistically (tried two phase screens counter rotating back to back)
  - Errors in source location on the calibration bench
  - Turbulence above maximum bench phase screen altitude (approx. 10km)
  - Extra latency (cant stream centroids through net as they arrive, need all of them)
    - Processing time (carmen) = 1.01 +/- 0.01 ms
    - Processing time (L&A) = 0.68 +/- 0.02 ms
    - cf. median τ_0 = 5 ms
- Despite these issues:
  - Within 5% Strehl ratio of L&A
  - Within 15nm WFE rms of L&A

Future work

- Train with on-sky data
- Predictive ANN
  - Train ANN with t=0 guide star slopes to predict t+dt slopes.
- GLAO
  - 'noise' from high altitude layers reduce performance of GLAO. Can we train an ANN to ignore/use this? 'tomographic GLAO'?

Conferences in Durham, 15th -19th Sept. 2014

- 'Adapting to the Atmosphere'

Atmospheric Profiling and Adaptive Optics conference, Durham 15th – 17th September, 2014

https://www.dur.ac.uk/cfai/sitecharacterisation/profconf/

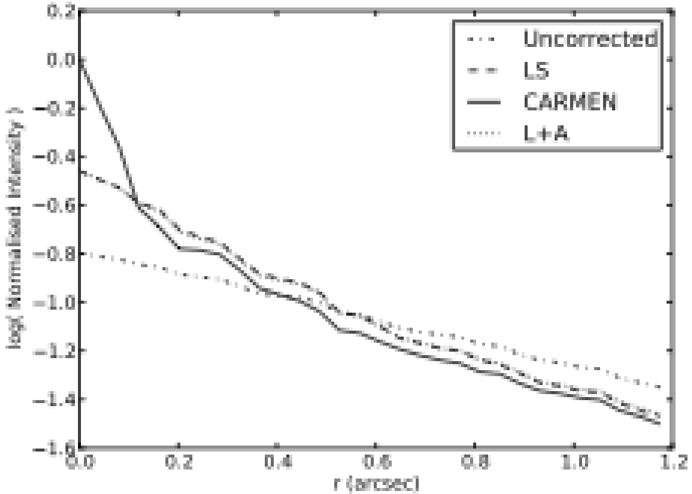
- AO Modelling and Simulation Workshop

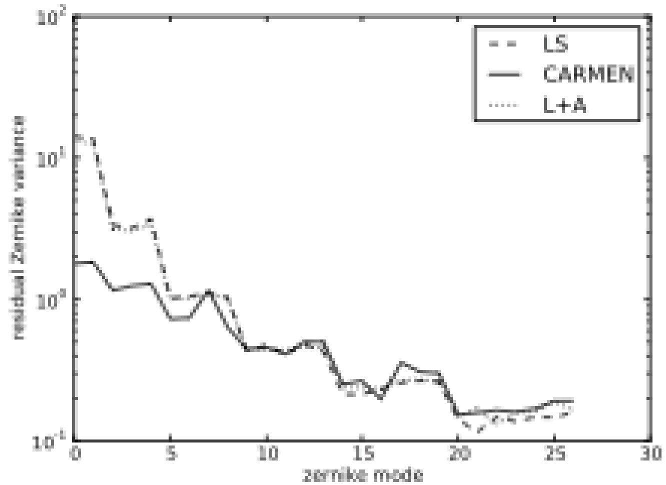
18th – 19th September, 2014

https://www.dur.ac.uk/cfai/adaptiveoptics/sim2014/

Any questions or comments?

Simulations with Shot and Read noise





- Trained with expected noise levels (different to Montera et al.)
- 100 photons per subaperture (m=11, D=4.2, n =7, npix = 20)
- 0.2 e^- readout noise