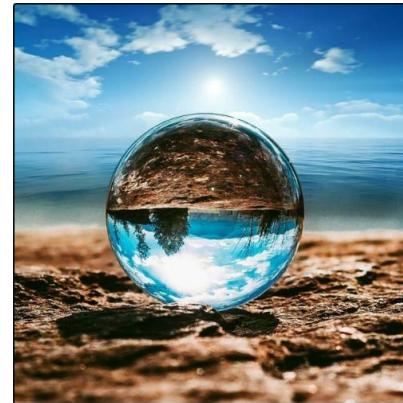


The Seer



Victor Madrid, Tate Harsch-Hudspeth, Evan Peelen
Sonoma State University, Department of Engineering Science
Advisor: Dr. Mohamed Salem
Industry Advisor: Joe Reid, Radio Hobbyist
Client: Cellular Providers, Cellular Application Developers

Evan: Evan.Peelen@comcast.net
Victor: vmad1999@gmail.com
Tate: tatehh411@gmail.com

<https://harschht.github.io/The-Seer/>

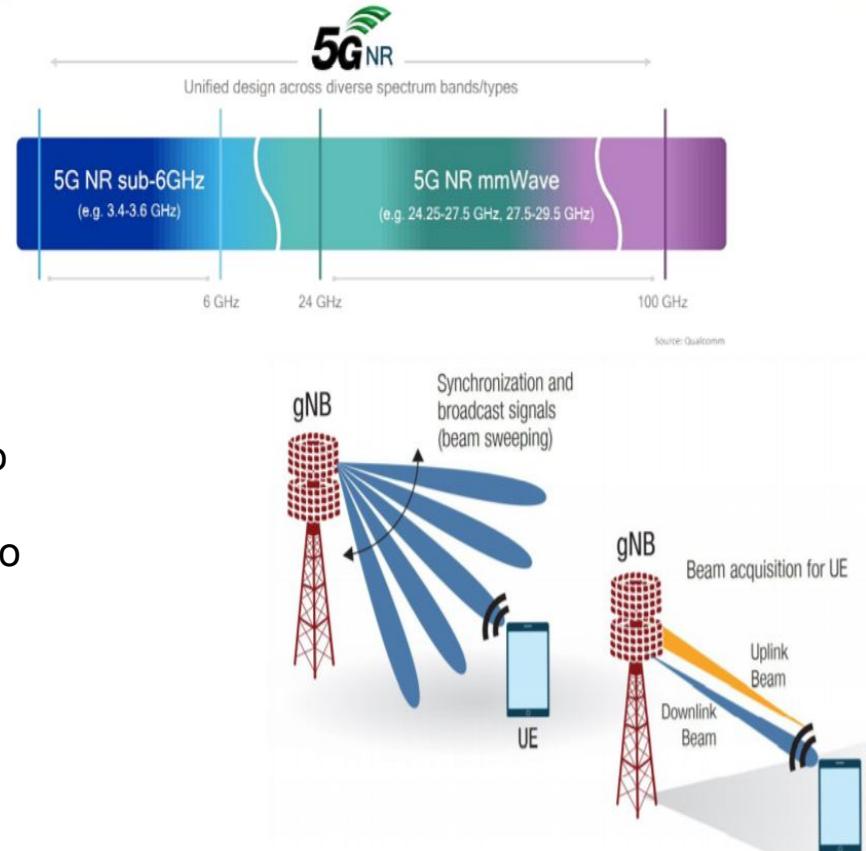
- Problem Statement
- Existing Solutions
- Marketing Requirements
- Engineering Requirements
- System Diagrams and Flowcharts
- Challenges & Risks
- Tests
- System Components & Budget
- Schedule



KORET
FOUNDATION

The Problem

- Methods commonly used to find DOA assume an isotropic environment and do not account for:
 - Multipath
 - Constructive/Destructive interference
 - Non-proportional propagation
 - Changes within the environment
- This can mean poor performance in complex environments where there are many contributors to the EM field
- With a non-proportional environment, the solution to this inverse problem can be extremely difficult, increasing the challenges of modeling the propagation medium
- DOA technology could benefit smart device applications that rely on precise location for their services (AGPS)



Existing Solutions

Other DOA Models

- MUSIC Algorithm (MULTiple SIgnal Classification)
- SAMV (iterative sparse asymptotic minimum variance)
- ESPRIT (estimation of signal parameters via rotational invariant techniques)
- Periodogram

Full Dimension MIMO (FDMIMO)

- Great at determining the best path for signal strength
- Cares little about the location the signal is coming from
- Great for cell signals, but not for applications relying on precise location services (AGPS)

Triangulation

- Uses timing advance and power levels of incoming signals to determine Tx location
- Communication between multiple base stations is required, leading to excess power use

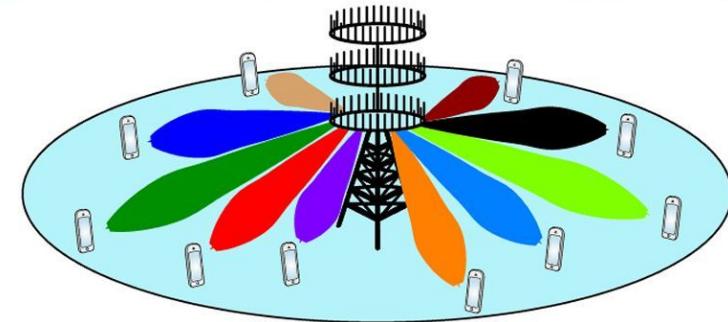


Figure: (above) FDMIMO

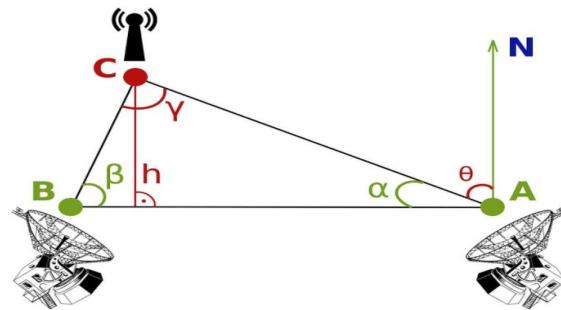
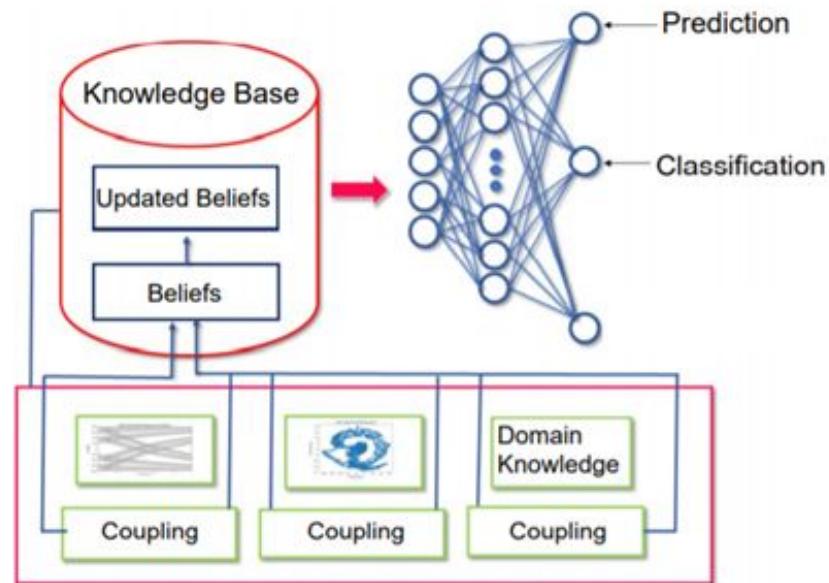


Figure: Triangulation



Proposed Solution

- Implement a system that uses Deep Learning to determine the direction-of-arrival of an incoming low-band 5G signal
 - Can work well indoors since the environment need not be isotropic
 - NN will develop a model that accounts for the complexity of the environment
- Sub-6 5G NR location estimation within the 600 MHz (n5) to 850 MHz (n71) band
- Proof of concept prototype will be able to solve the aforementioned problem by mapping the EM spectrum within the testing environment through training





Marketing Requirements

- MR-1) The system must streamline the process of the Rx determining the direction-of-arrival of the incoming signal (less time and less energy than triangulation)
- MR-2) The system must be able to determine direction-of-arrival within an acceptable range
- MR-3) The system can be modified for other environments through training of the neural network
- MR-4) The system must be able to handle noise up to a certain threshold
- MR-5) The system must be able to understand and work with low-band 5G signals
- MR-6) The project should have a interface where the user can see data clearly
- MR-7) The system must be inexpensive enough for mass production

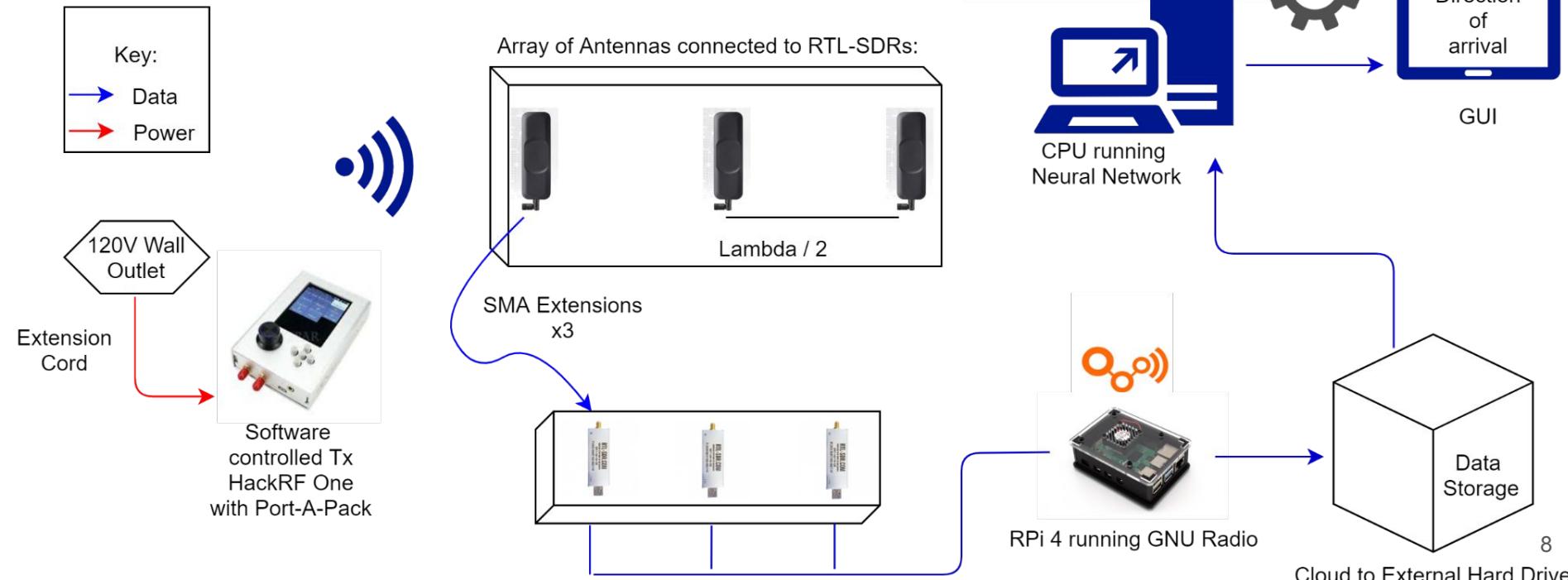


Engineering Requirements

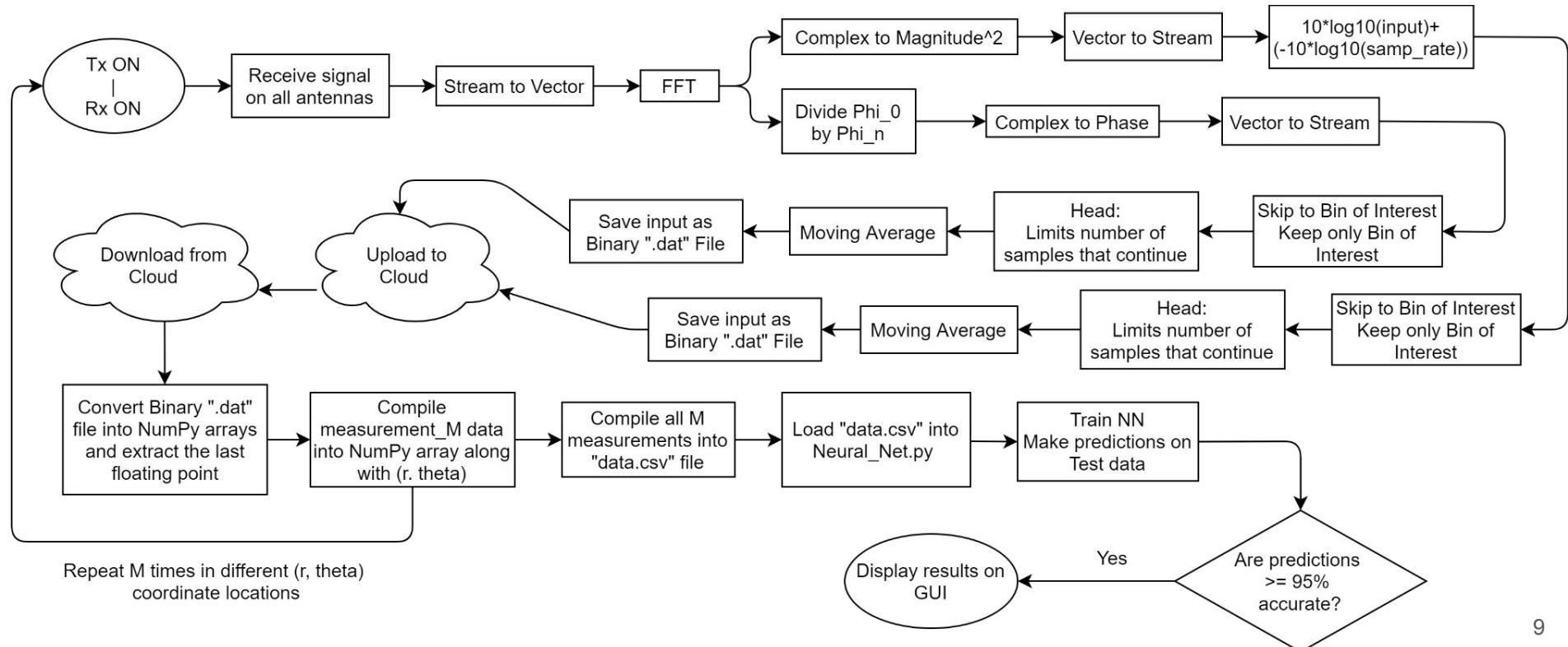
- ER-1) The network must be able to guess the direction-of-arrival of the transmitted signal with an accuracy level of 90% or greater
- ER-2) The system must work for 5G signals transmitted within a radial distance of at least 6 meters
- ER-3) The system must be able to come up with a valid model for any environment it is trained in
- ER-4) The system must be accurate in the presence of ≤ -40 dB of noise
- ER-5) The system must be able to work with frequencies in the 600-850 MHz band
- ER-6) There must exist a GUI that displays real and accurate data within 60 seconds
- ER-7) Prototype must be less than \$600

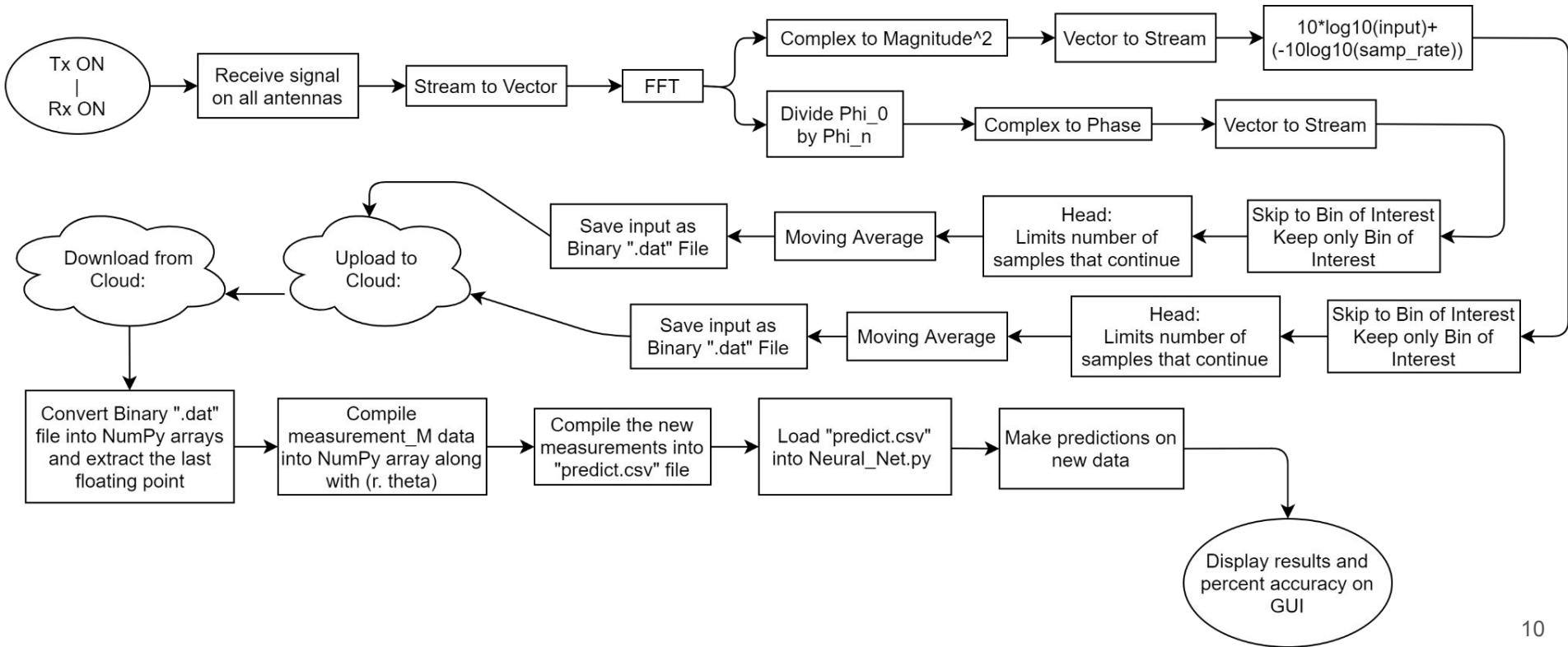


System Diagram



Software Flowchart - Training Stage







Anticipated Risks

STATE UNIVERSITY

Hardware:

- Uncontrolled RF testing environment will make testing hardware difficult
- Limited USB ports, so USB hub will be necessary
- Antennas do not come with SMA extensions for antenna spacing, low loss coax cables will be required
- Signal degradation due to multiple connections and transmission lines, or noise generated by loose connections on the SDR clocks could lead to offsets in phase data

Software:

- There is a chance we will overfit our data, resulting in less accurate predictions
- We may overlook environmental factors and fail to create a complex enough model to make accurate predictions
- If our data extraction flowgraph is not representative of the true RF power and relative phase difference of the received signal, our N.N. will not be able to make accurate predictions



Challenges

STATE UNIVERSITY

Hardware:

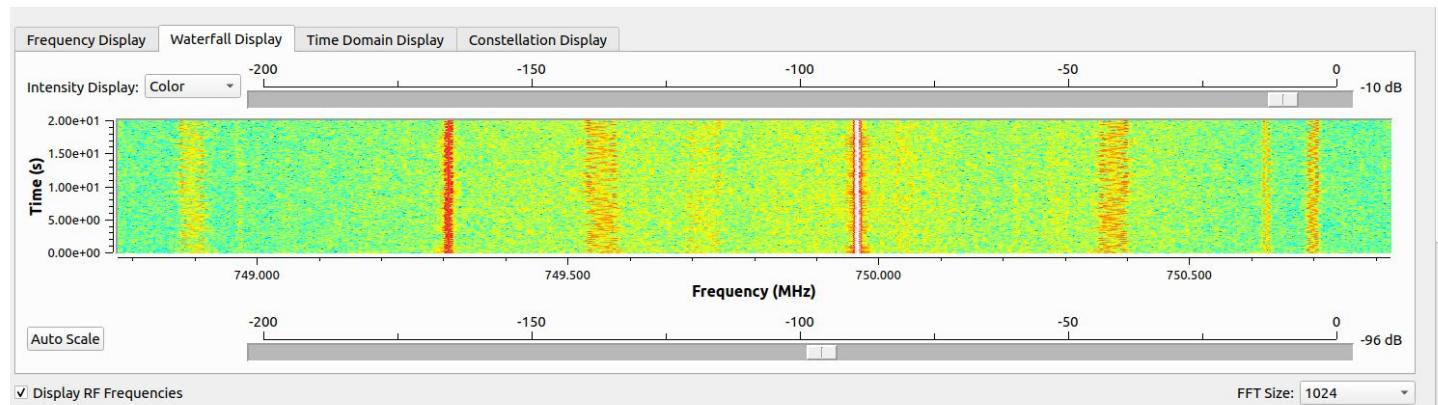
- Limited lab equipment access due to COVID-19 impact
- Delay in funding availability due to COVID-19 impact
- CPUs built before 2011 do not support Advanced Vector Extensions (AVX)
- The SDRs are power hungry devices, our CPU hardware was unable to drive all 5 SDRs, and our antenna array design needed to be adjusted
- Phase drift (time delay) at the Rx caused by synchronizing the SDR clocks disrupted the sampling rate, affecting the credibility of our phase data

Software:

- Most of the documentation regarding Keras and GNU Radio is geared towards specific applications, requiring us to build our own models
- GNU radio does not contain all libraries in default download, so additional libraries needed to be found and installed
- Open source software documentation was limited, requiring additional testing and research

Hardware (continued):

- Third party HackRF One had a faulty mixer which caused the center transmitting frequency to be offset by 40 kHz at 750 MHz (offset grew for higher frequencies and decreased for lower frequencies)
- The RTL-SDR used for testing was damaged due to prolonged exposure to high input power resulting in unwanted frequency shift keying introducing uncertainties into our data analysis





Test Overview

STATE UNIVERSITY

Tests:

Test 1A - Proving that an AM signal is scalable by distance, ER - 1,2,3,4

Test 1B - Proving that an FM signal is scalable by distance, ER - 1,2,3,4

Test 1C - Finding the environmental path loss exponent, ER - 1,2,3,4

Test 2A - Neural Network assessment using Sim. 01 (Python), ER - 1,3

Test 2B - Neural Network assessment using Sim. 02 (MATLAB), ER - 1,3,4

Test 3 - Test coax cables and find S11 parameters of antennas, ER - 1,2,4,5

Test 4 - Antenna Array clock sync. time delay analysis, ER - 1,3,5

Test 5 - Flowgraph test with 5 RTL-SDRs, ER - 1,2,3,4,5

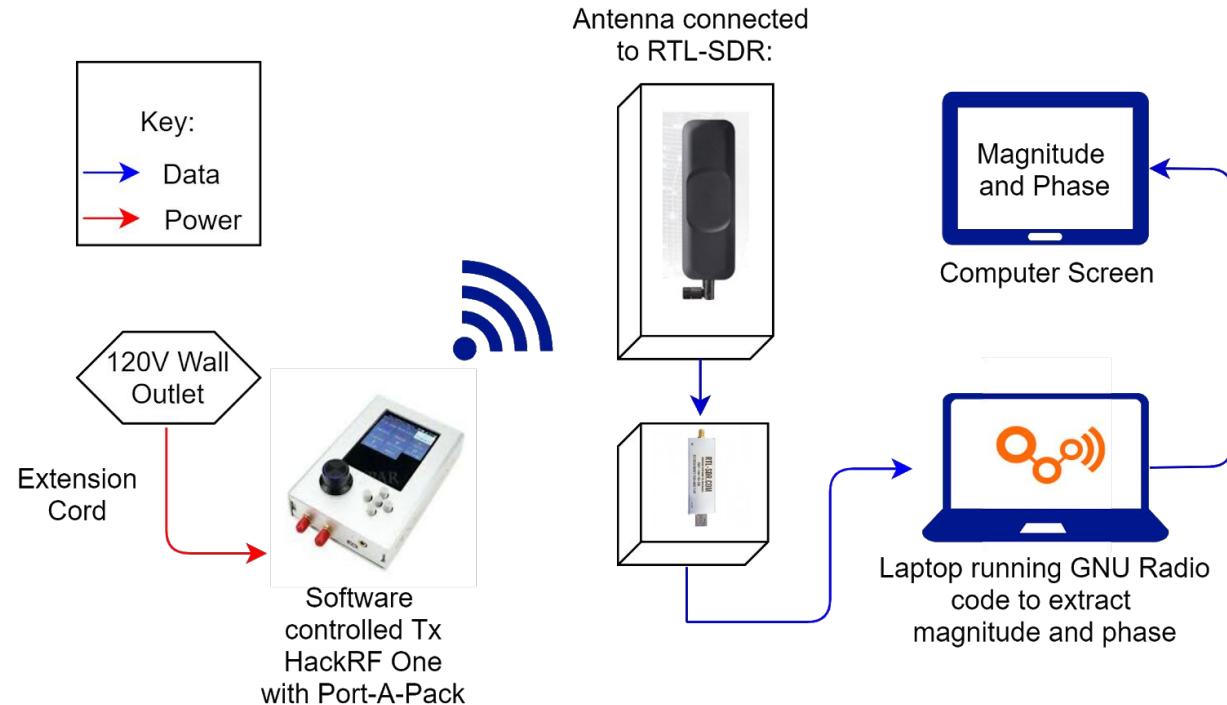
Test 6A - Antenna Array test environment power calibration, ER - 1,2,3,4,5

Test 6B - Antenna Array test environment phase calibration, ER - 1,2,3,4,5

Test 7 - GUI display test and result display-time analysis, ER - 1,6

Test 1C- Objective and Setup

Our goal is to verify that our hardware and testing environment are representative of the team's expectations when transmitting and receiving RF waves in an urban setting





Test 1C

Test 1C - Finding the Urban Outdoor Path Loss, ER- 1,2,3,4

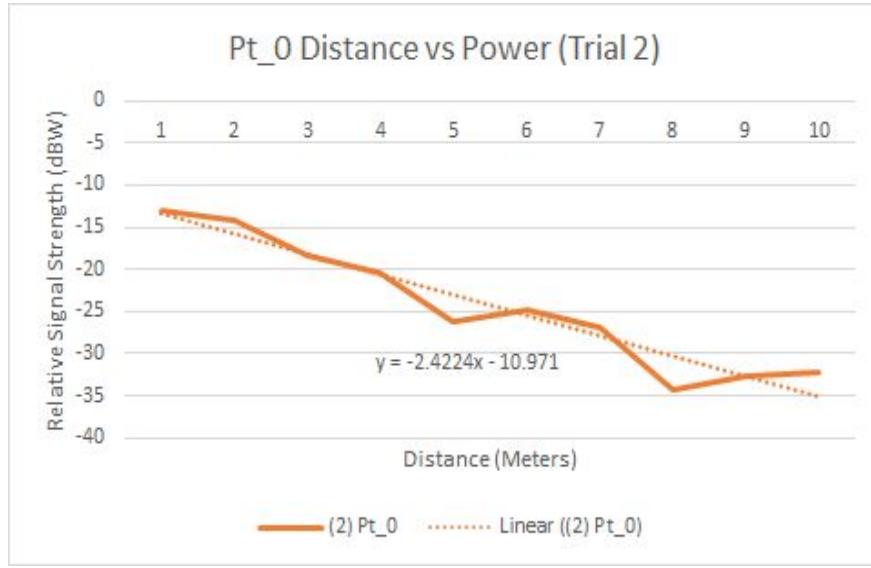
Tests and Unit/System to be Tested	Objective	Expectation
Transmitter and Receiver	Find the “free space” attenuation constant using one SDR and the Port-A-Pack. Since the team cannot access a true “free space” environment, this test would be conducted in an outdoor-urban environment.	We expect the urban outdoor path loss, alpha, to be between 2 and 3.5 dBW/m. We expect there to be a relative level of noise due to the urban environment. Note: This test was a success.

Test 1C Results

STATE UNIVERSITY

Typical Path Loss Exponents for Common Environmental Settings

Environment	Path Loss Exponent (dBW/m)
Free Space	2.0
Urban Area Cellular Radio	2.7 to 3.5
Outdoor-Urban (The Seer Test Environment)	2.0 to 3.5
In Building Line-of-Sight	1.6 to 1.8
Obstructed in Building	4.0 to 6.0



The Seer Path Loss Exponent:
2.422 (+/-) 0.19 dBW/m

Path Loss Exponent (Attenuation Constant):

- Expected path loss exponent in an outdoor-urban environment is 2-3.5 dBW/m
- The team would calculate the path loss exponent in their outdoor-urban environment to be $2.422 (+/-) 0.19$ dBW/m
- Calculated path loss exponent meets expectation for test environment
- Successfully verified that our hardware and testing environment are representative of the team's expectations when transmitting and receiving RF waves in an outdoor-urban setting



STATE UNIVERSITY

Our aim is to use various RF propagation models to assess our Keras models ability to handle relevant data.

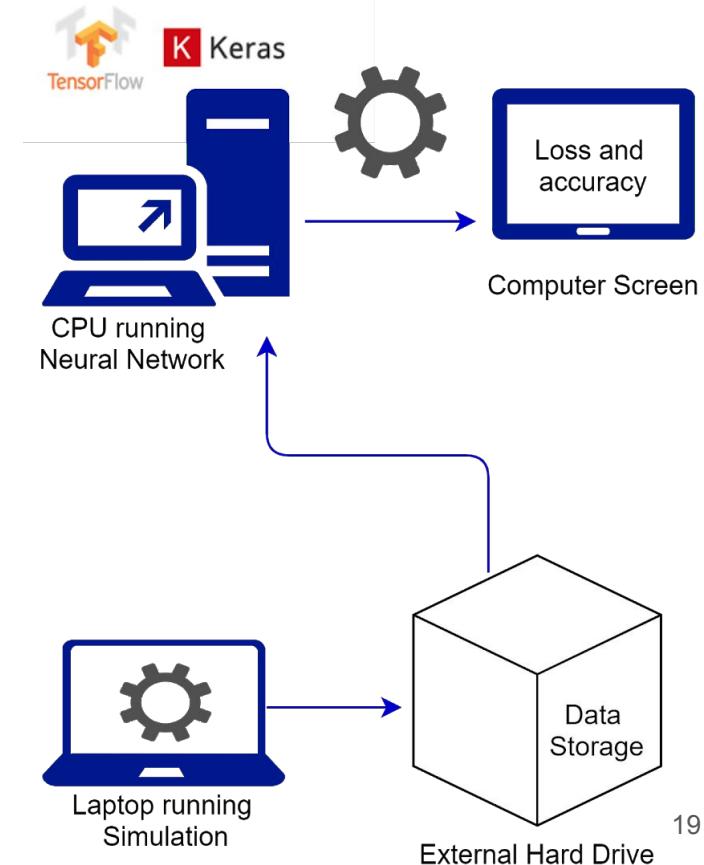
Friis equation: (ideal case)

$$\frac{P_r}{P_t} = D_t D_r \left(\frac{\lambda}{4\pi d} \right)^2 V = \sqrt{P \cdot Z_0} \quad \phi = \omega * (l/c)$$

JTC model: (indoors w/ added noise)

$$L_{Total} = A + B \log_{10}(d) + L_f(n) + X_\sigma$$

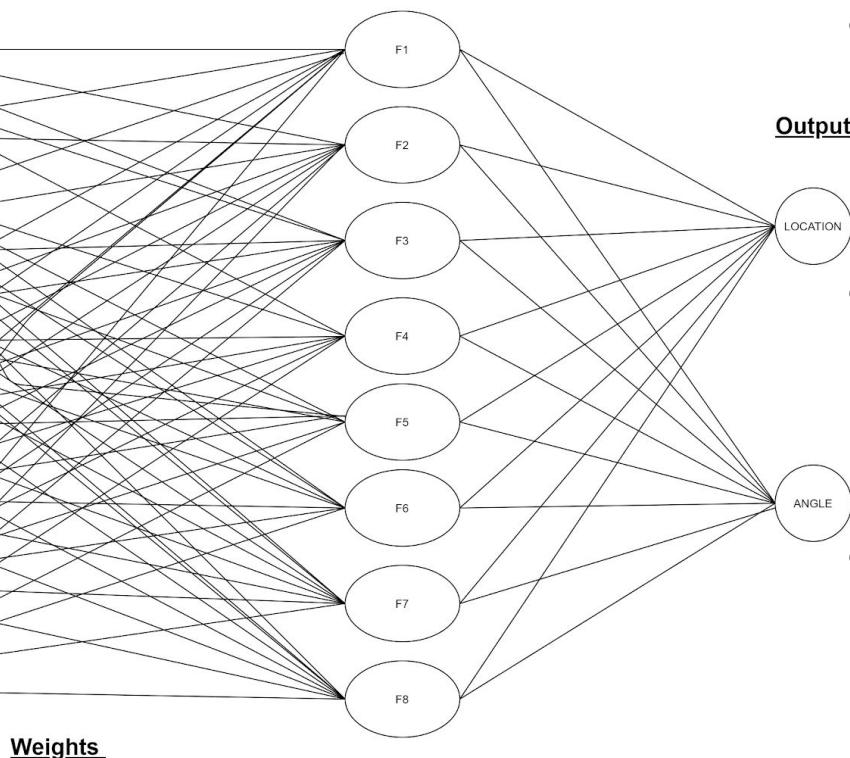
In order to make the JTC simulation more representative of our final testing environment we factored in a white gaussian noise level corresponding to the amount of noise present in previous tests. The JTC model also accounts for walls and obstructions.





Test 2

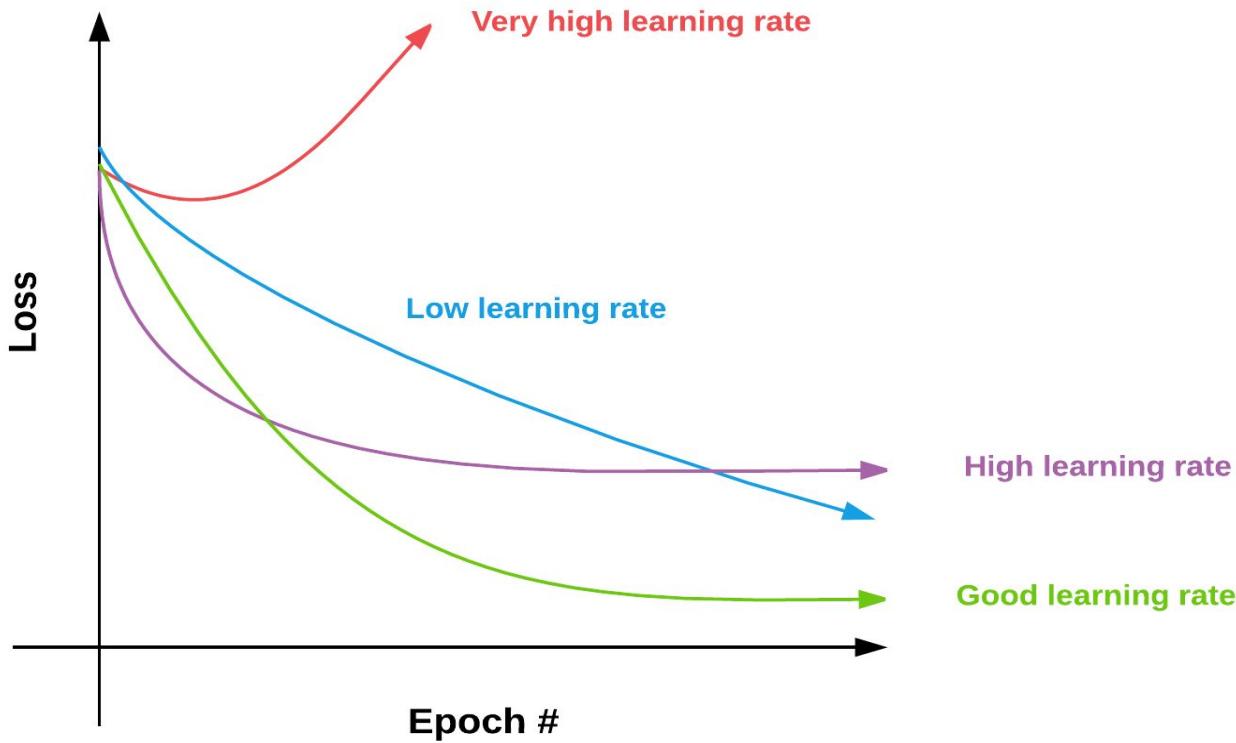
Test	Unit/System to be Tested	Objective	Expectation
Test 2A - Neural Network assessment using Simulation 1 (Python) ER- 1,3	We are analyzing our Keras model on our original 5AA design. Our model attempts to solve our multiple output regression problem. Test 2A handles the ideal free space scenario.	Use the Friis transmission equation free space path loss RF propagation model to asses our Keras models ability to handle relevant data.	We expect our model to handle this RF propagation model well, producing a high loss that falls relative to the number of epochs run. Note: This test was a success.
Test 2B - Neural Network assessment using Simulation 2 (MATLAB) ER- 1,3,4	We are analyzing our Keras model on our original 5AA design. Our model attempts to solve our multiple output regression problem. Test 2B handles a more complex scenario representative of our final testing environment.	Use the JTC transmission equation indoors environment path loss RF propagation model to asses our Keras models ability to handle relevant data.	Since the propagation model is slightly more complex, and we have added gaussian noise to the signals, we expect to see a more drawn out learning curve, needing more epochs to asses the problem. Note: This test was a success.



What is a Neural Network?

- Each neuron (node in the diagram):
 - Computes weighted sum of all inputs multiplied by each weight
 - In a network of n inputs:
$$\text{node}_j = \sum_{k=1:n} (x_k * w_{kj})$$
- A NN consists of a specified number of neurons creating a transfer function:
 - Stochastic gradient descent is used to find the model resulting in global minimum error
- An optimization function determines how to update the weights in the neurons based off of the value of a loss function

Expectations - Neural Network



Epoch

How many times the NN moves through the entire dataset.

Loss

A scalar value that represents how incorrect the models predictions are.

Batch Size

How many samples the NN processes before updating the weights



Analyzing the Results:

- Using 25 predictions, our model produced an average difference:

$$\frac{\sum(\text{ABS}(y_{\text{predicted}} - y_{\text{collected}}))}{25}$$

(The predictability of the NN, utilizing test 2A data)

of 0.38 meters for the R value, and 4.42 degrees for Theta, with a standard deviation of 0.42 for R and 4.43 for Theta

(The predictability of the NN, utilizing test 2B data)

of 0.88 meters for the R value, and 11.9 degrees for Theta with a standard deviation of 0.64 for R and 6.55 for Theta

2A & 2B Results + Conclusion

Conclusion:

- Friis propagation model performed best after training over 1,150 epochs while the JTC w/ added noise performed better after 255 epochs. The JTC model had twice as many weights per hidden layer compared to the Friis.
- The added complexity of the JTC model provided evidence that our physical system needed a more complex model (added layers or weights) in order to capture the complexity of the environment.
- Test 2 results verify that our neural network is capable of achieving the accuracy stated in our engineering requirements.

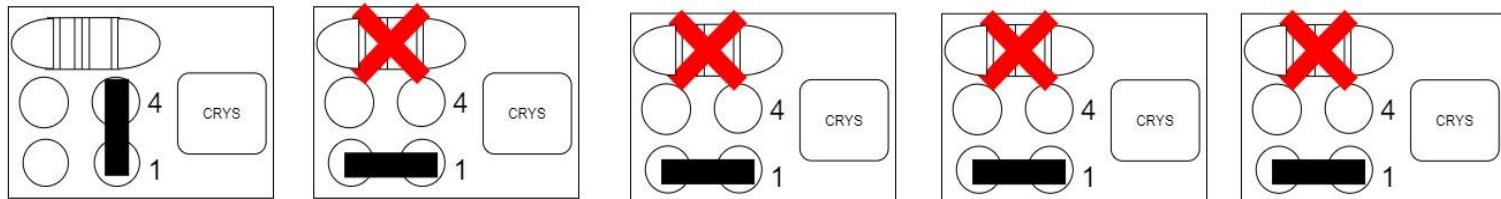
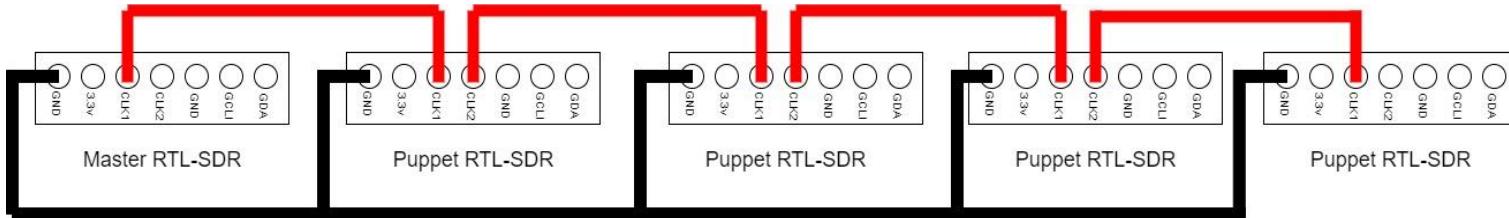
Test 4 - Antenna Array Clock Sync

Test 4 - Antenna Array clock sync. time delay analysis, ER - 1,3,5

Objective:

- Daisy chain the RTL-SDRs together
 - Ensure that there is little to no latency in the clock period between them

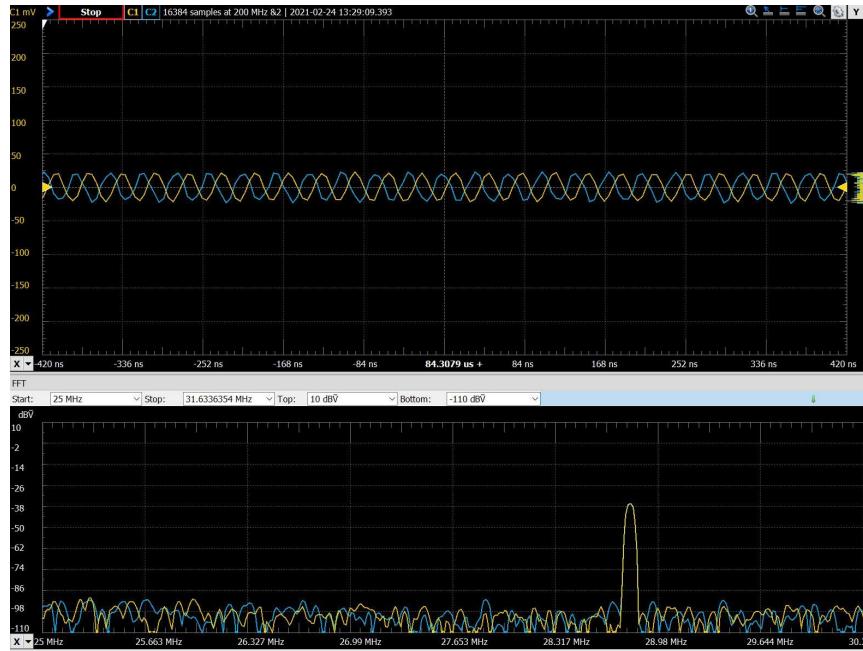
Setup:



Test 4 - Results & Conclusion

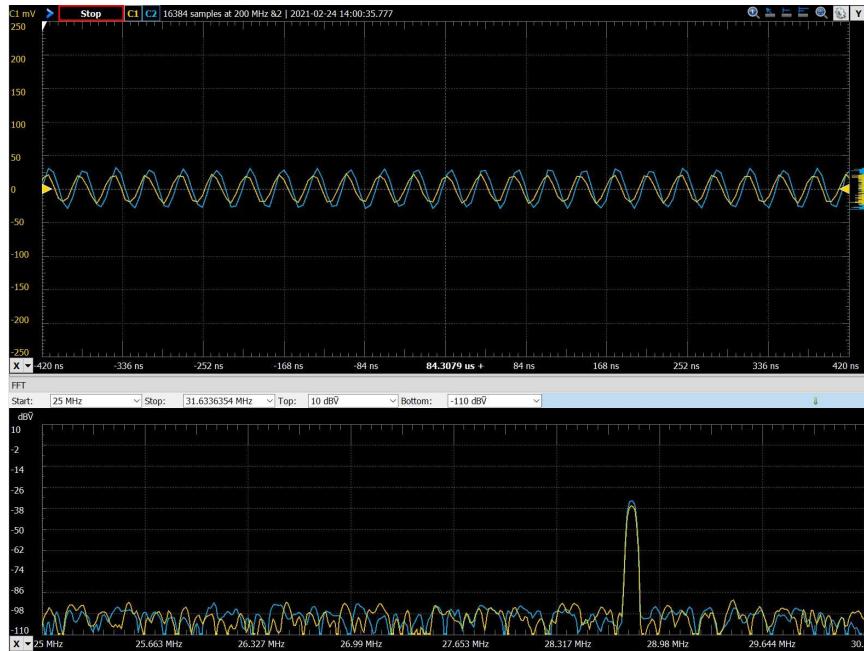
- There is a 0.019 ns shift in period between the master and the first puppet receiver

Master clock (yellow) vs. Puppet clock #1 (blue)



- There is a 0.012 ns shift in period between the master and the last puppet receiver.

Master clock (yellow) vs. Puppet clock #4 (blue)



Test 5 - Flowgraph Test: 5 RTL-SDRs

Tests and Unit/System to be Tested	Objective	Expectation
Receiving Antenna Array and 5 AA GNU Radio Flowgraph	Run the GNU Radio Flowgraph using the 5 antenna array and verify that the values recorded are consistent with the team's expectations.	<p>We expect there to be issues with the amount of output power since the V3 RTL-SDRs consume 275mA of current while a USB 3.0 port on the RPi 4 can supply 1.2A to peripheral devices. To combat this issue we used a USB Hub that is capable of drawing 1.2A from a standard 120V wall outlet.</p> <p>Note: This test was a failure, causing the team to reduce the antenna array from 5 SDRs to 3 SDRs.</p>

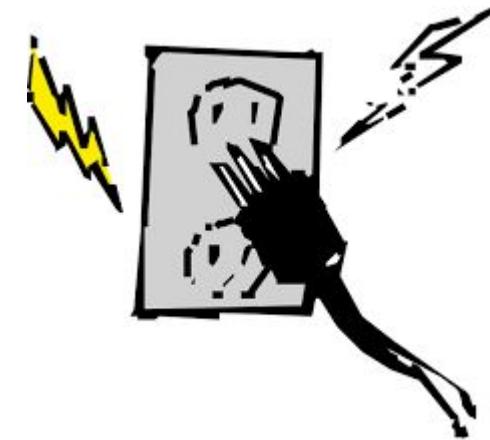
Setup:

5 Antenna Array



Test 5 - Results & Conclusion

- Since 5 SDRs together would require 1.35A of current, we added a USB hub capable of supplying 1.2A from a wall outlet
- Unfortunately, this did not solve our power issue and the RPi4 was not able to run all 5 SDRs simultaneously
- Flowgraph would run for a few seconds and then the SDRs would disconnect and become unrecognizable by the SoC
- We tried running them at the lowest allowable sampling rate of 256KSps to no avail
- Also tried running the flowgraph using a PC with an intel i7 core processor and encountered the same scenario
- When the 5 SDRs are run simultaneously without enough power to drive them, they engage a failsafe that disconnects them from the computer and keeps them from reconnecting until they have been powered off and rebooted
- This is to protect the PCB and memory on board the SDRs



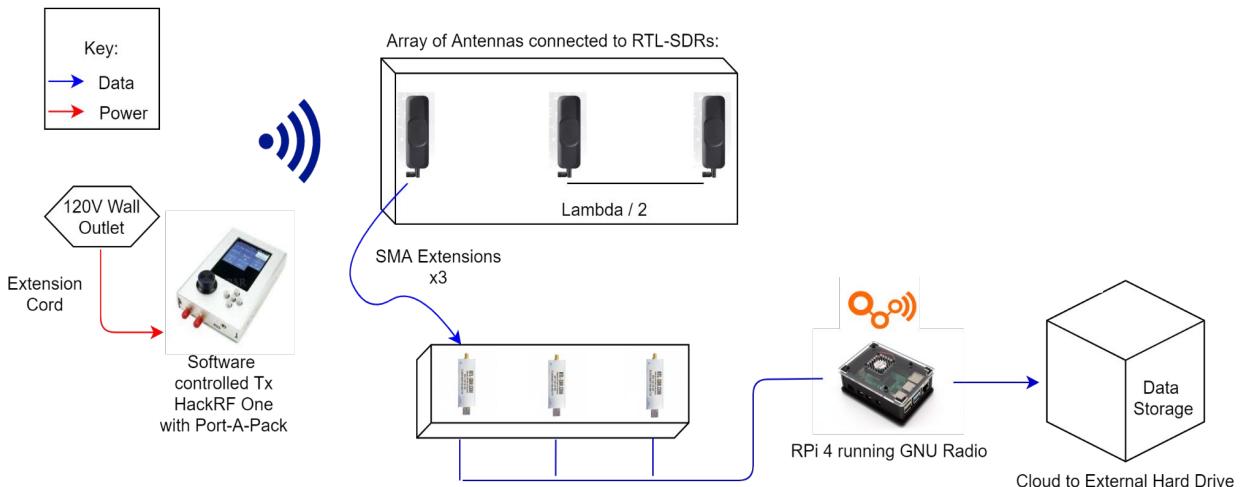


Test 6 - Antenna Array Calibration

Test	Unit/System to be Tested	Objective	Expectation
Test 6A - Antenna Array Power Calibration ER- 5, 2	We are testing our reduced 3 antenna array and data extraction flowgraph to ensure that the extracted magnitudes are consistent with the team's expectations.	Our goal for this test is to confirm that our antenna array power data will be useful to the neural network.	<p>We expect to find an attenuation constant similar to the results of Test 1C. The extracted power should be scalable by distance. We are looking for any distortions or uncertainties that would suggest that there are problems with the antenna array construction.</p> <p>Note: This test was a success.</p>
Test 6B - Antenna Array Phase Calibration ER- 5, 2	We are testing our reduced 3 antenna array and data extraction flowgraph to ensure that the extracted relative phase differences are consistent with the team's expectations.	Our goal for this test is to confirm that our antenna array relative phase difference data will be useful to the neural network.	<p>The results from Test 4 verified that the clocks were synchronized, so there will not be any constant phase difference between receivers. If the phase changes with time, however, then the phase data will be useless in determining position.</p> <p>Note: This test was a failure, causing the team to drop the relative phase difference from the N.N. inputs.</p>



Test 6 - Setup



STATE UNIVERSITY

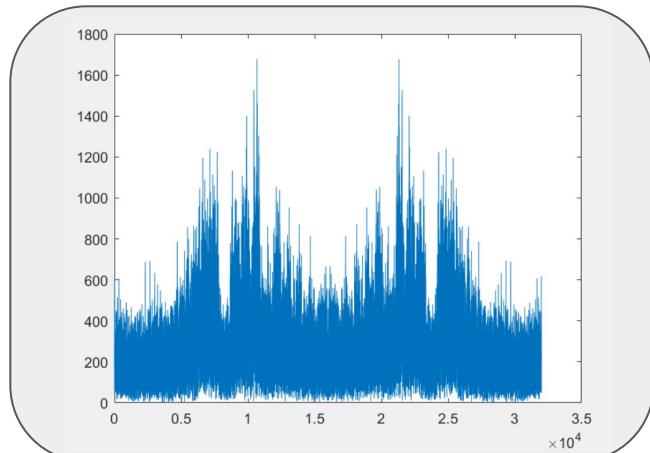
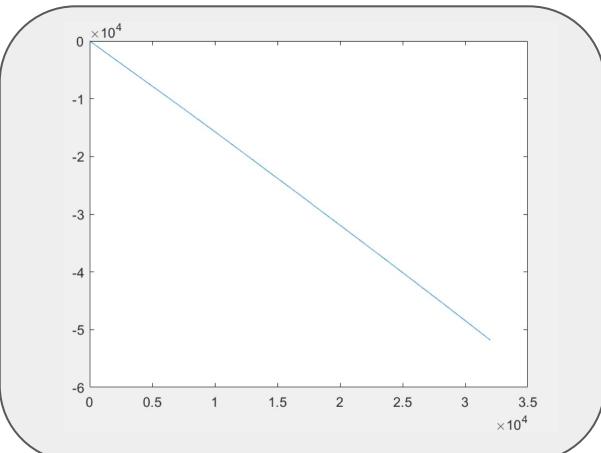
Test 6A:

- The extracted magnitude corresponded to the distance between the Tx and Rx, verifying that our power data would be useful to the neural network

Test 6B:

- The quality of the SDR clock used as the master, along with an increased temperature of the SDR PCBs with the heat sinks removed caused a delay between the sampling of each SDR, accumulating a phase drift that only grew as time passed. This rendered the phase data useless, as the target position information was distorted by this effect

Antennas (1-3)	Path Loss Exponent (dBW/m)
Antenna 1	4.005 +/- 0.7087
Antenna 2 (Reference)	2.174 +/- 0.4885
Antenna 3	2.260 +/- 0.5371



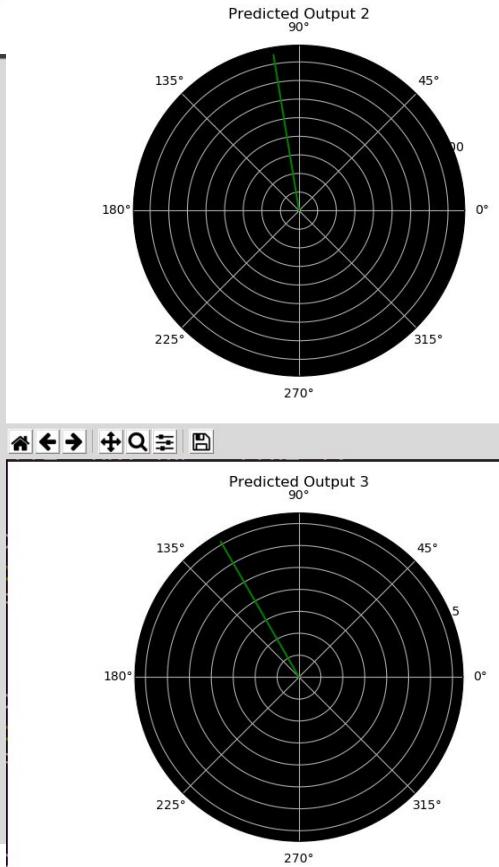
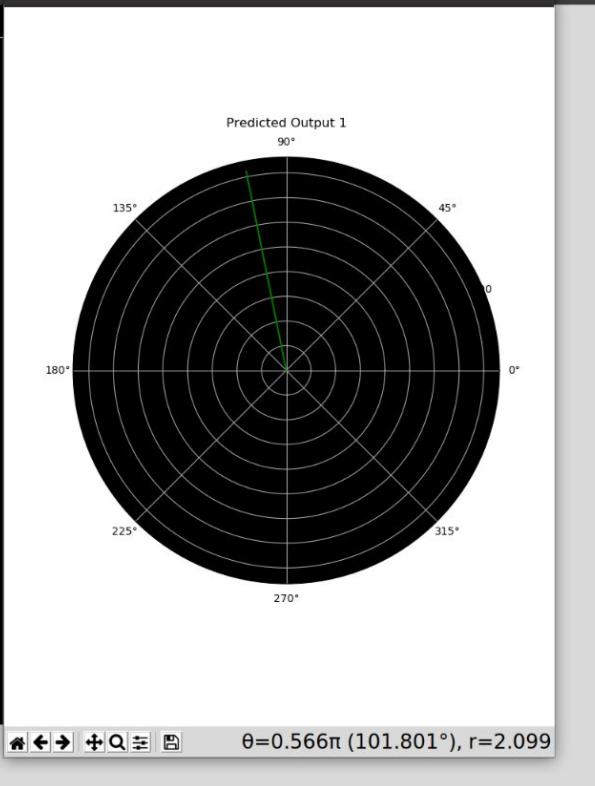


SONOMA

STATE UNIVERSITY

System Results

5G DOA		
THE SEER		
Inputs		Predicted Outputs
Power [dBm]		Distance Angle
-34.0	-35.0	2.1m 100.0°
-33.0	-38.0	2.1m 99.0°
-36.0	-38.0	3.6m 120.0°
-38.0	-35.0	2.8m 94.0°
-40.0	-38.0	3.4m 82.0°
-39.0	-35.0	2.5m 79.0°
-47.0	-34.0	3.6m 79.0°
-43.0	-40.0	3.3m 81.0°
-36.0	-37.0	3.3m 110.0°
-38.0	-40.0	3.3m 85.0°
Avg. Distance Diff.	0.65m	Avg. Theta Diff. 22.0°
		System Accuracy 90%





System Results

- With the reduction from 5 antennas to 3, the neural network lost a significant amount of data, reducing the number of inputs from 8 to 5, impacting our systems ability to accurately estimate the DOA
- With the removal of the relative phase difference, the neural network lost even more data, reducing the number of inputs from 5 to 3
- Using this tensor made up of the 3 magnitudes, our system was able to predict the D.O.A. with a system accuracy of 90%
- Our system accuracy was based on the percentage of validation samples that passed our teams predetermined metric:
 - R difference < 1 meter
 - Theta difference < 45 degrees
- This is impressive and speaks to the ability of our deep learning algorithm to build a model representative of our complex inverse problem

5G DOA			THE SEER	
Inputs			Predicted Outputs	
Power [dBW]			Distance	Angle
-34.0	-35.0	-34.0	2.1m	100.0°
-33.0	-38.0	-33.0	2.1m	99.0°
-36.0	-38.0	-41.0	3.6m	120.0°
-38.0	-35.0	-36.0	2.8m	94.0°
-40.0	-38.0	-37.0	3.4m	82.0°
-39.0	-35.0	-33.0	2.5m	79.0°
-47.0	-34.0	-31.0	3.6m	79.0°
-43.0	-40.0	-40.0	3.3m	81.0°
-36.0	-37.0	-39.0	3.3m	110.0°
-38.0	-40.0	-37.0	3.3m	85.0°
Avg. Distance Diff.			Avg. Theta Diff.	System Accuracy
0.65m			22.0°	90%

Show Graph

Close



Prototype Cost

Item	Description	Price	Quantity	Cost
RPi 4 (8GB RAM)	System on Chip (SoC) used for processing the received signals through GNU Radio	\$104.99	1	\$104.99
32GB MicroSD Card	32GB MicroSD Card	\$8.25	1	\$8.25
RTL-SDRs	Software defined radio receiver capable of tuning frequencies of 500 kHz to 1.7 GHz	\$26.59	5	\$132.95
Antennas	TG.35.8113 Apex III Wideband 5G/4G Dipole Terminal Antenna with added 450 MHz band	\$16.46	5	\$82.3
Coax Cables	1M low loss SMA extension RF transmission lines	\$9.99	5	\$49.95
Software: Keras, OsmoSDR	Deep Learning Python library that uses TensorFlow2 as a backend Library of GNU Radio blocks used to run SDR Rx and Tx	FREE	NA	(open source)
PVC Pipe & Adhesive	Used for the antenna array structure	\$9.99	1	\$9.99
USB Hub	USB Hub with 1.2A output power from wall outlet	\$16.80	1	\$16.80
Total				\$405.23



Schedule

GANTT CHART AND PROJECT SCHEDULE

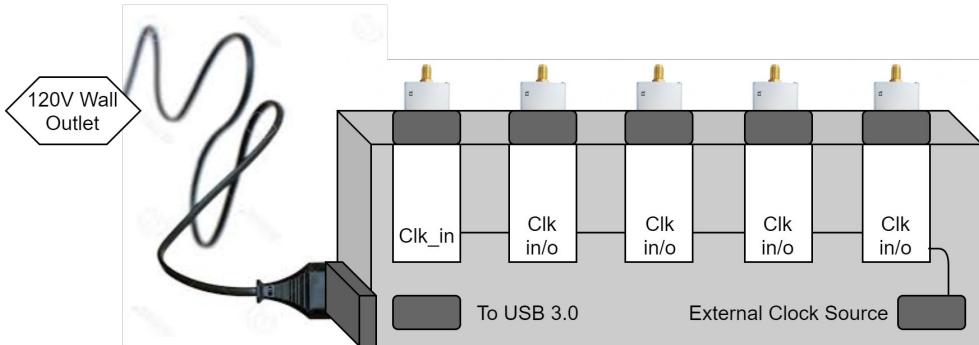




Future of The Seer

- Design external power supply/USB hub that can power 5 RTL-SDRs (V3 or V4) with 275 mA of current each, allowing bidirectional data flow to and from USB 3.0 port. In the design, include a synchronization of SDR clocks via highly accurate, temperature steady external clock source.
- Ensure that SDRs are truly synchronized, forming a coherent Rx without any phase drift
- Use external power supply to successfully run our 5-antenna flowgraph
- Run over 500 measurements worth of data through the 5-antenna neural network, test on over 100 validation measurements, accomplishing >95% system accuracy up to 12m from the Rx - [ER 1,2]

- We created a GitHub repository with all of project codes for both 3AA and 5AA in hopes that a future group of engineering students will continue the work we have started.





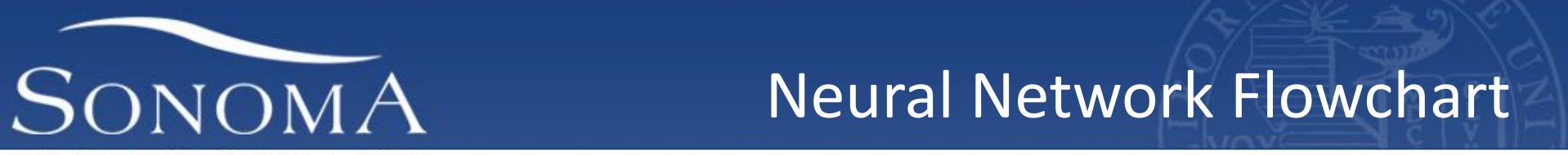
Thank You!
Questions
or
Comments?



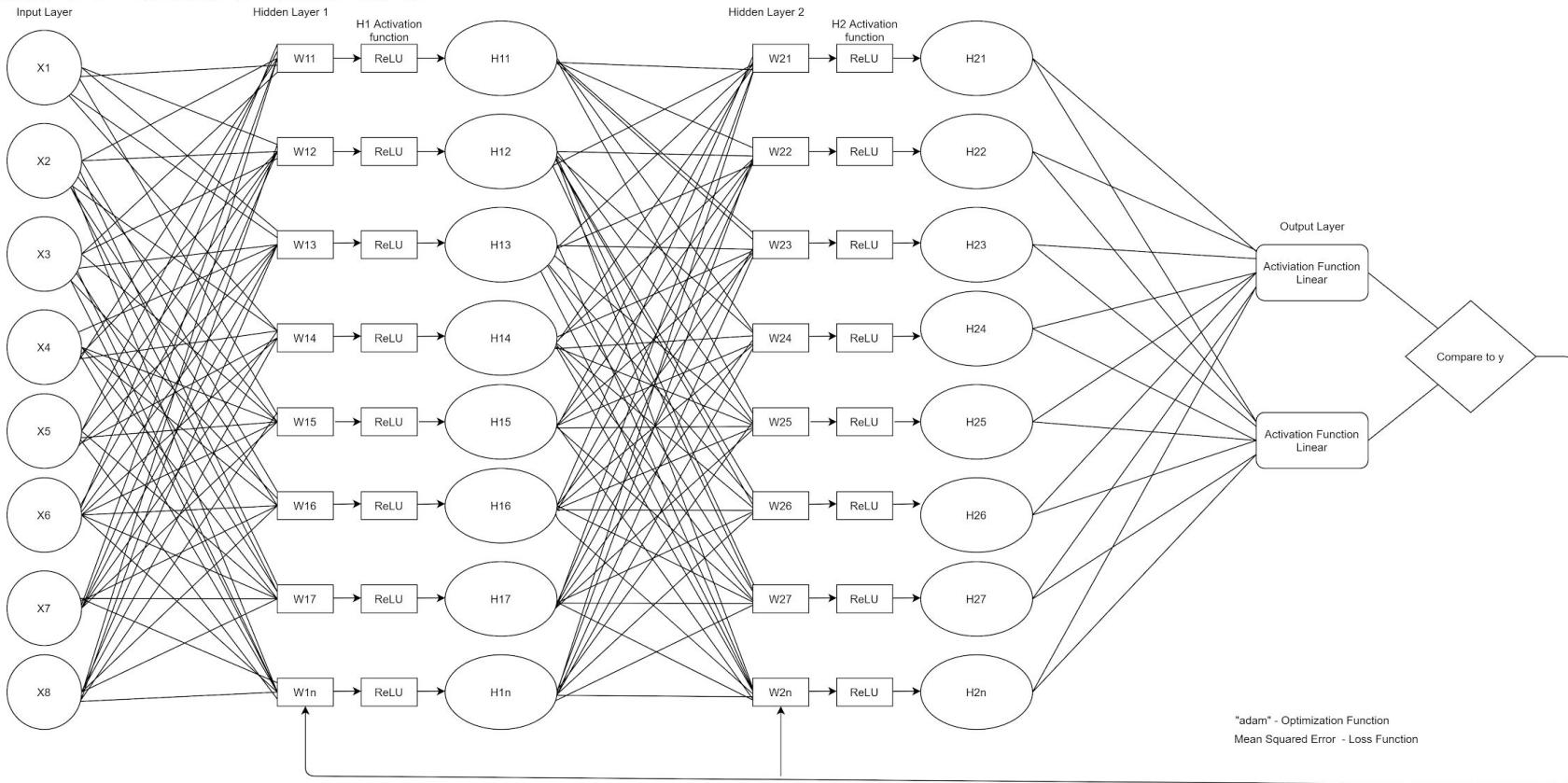


Image References:

- <https://www.amazon.com/Duosuny-Artificial-Crystal-Decoration-3-15inch/dp/B077FPV52W>
- https://line.17qq.com/articles/ffwqfnpy_p3.html
- <https://www.everythingrf.com/community/what-is-5g-new-radio>
- <https://www.4g-lte.net/technology/massive-mimo-technology/>
- https://en.wikipedia.org/wiki/Direction_finding
- <https://becominghuman.ai/artificial-neural-network-summary-and-tools-fdbd8f7fcb60>
- <http://www.clker.com/clipart-10821.html>
- https://onepiecetreasurecruiseglobal.fandom.com/wiki/Monkey_D._Luffy_Straw_Hat-Heart_Pirate_Alliance

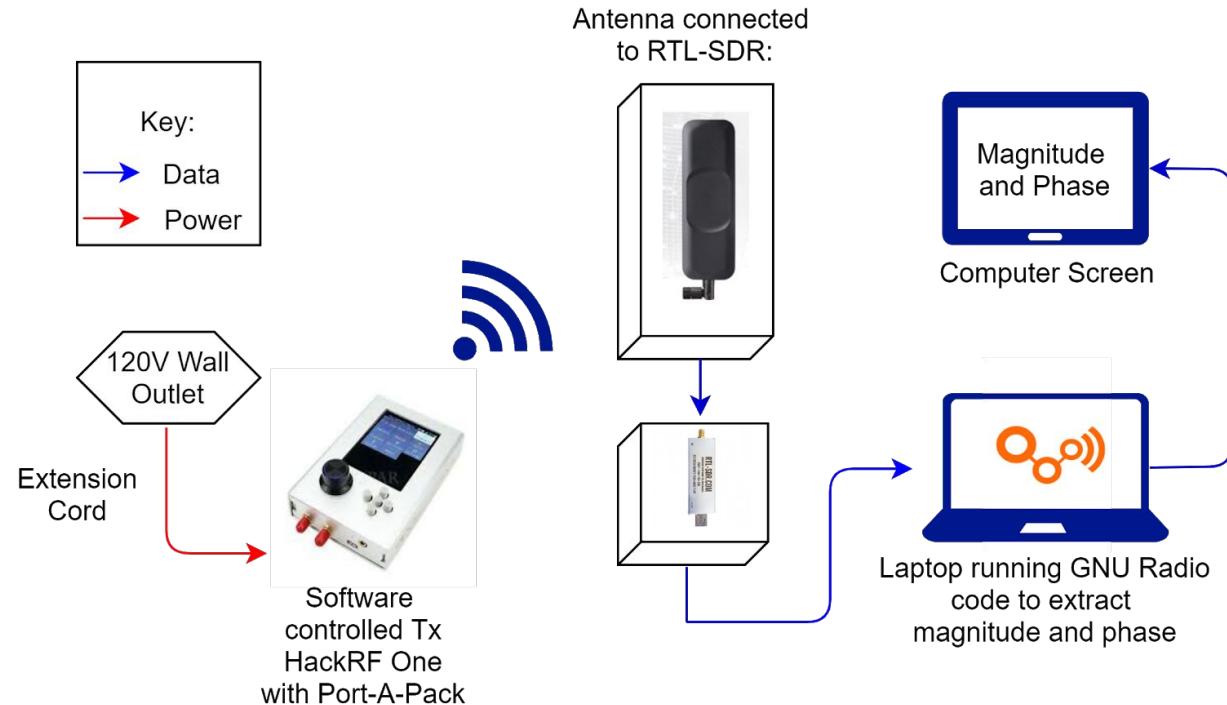


Neural Network Flowchart

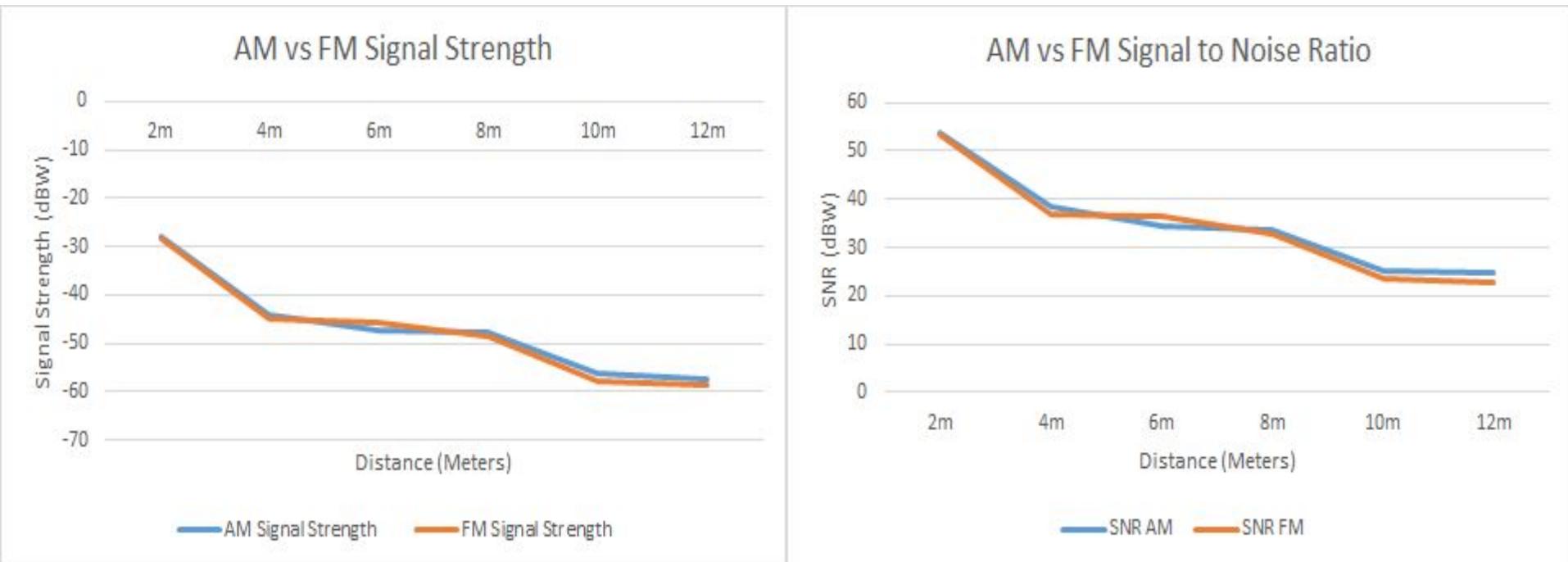


Test 1A & 1B - Objective and Setup

Our aim is to find quantities that are scalable by distance, to certify that our data will be useful to the Neural Network.



Comparing Tests 1A & 1B



Signal Strength (dBW)

- Both AM and FM signal strengths were scalable by distance and decreased in strength as the distance increased.
- Signal strengths for both AM and FM transmitted signals were comparable at each tested distance.

SNR

- The amount of signal power for both transmitted signals relative to the amount of noise in the testing environment were nearly identical, decreasing with increased distance.

Team Choice

- Team will use frequency modulation as the transmission modulation scheme, being highly comparable and nearly identical to our AM test results. We will use the HackRF One and the Port-A-Pack to Tx a 1kHz sine wave using FM.



Test 3 - S11 Return Loss

Return Loss:

- S11 represents how much power is reflected from the antenna, and gives us valuable information about how the antenna responds under certain frequencies.
- The maximum acceptable value of S11 for an interconnected structure is about -10 dB: If S11 is smaller than -10 dB, we will not see the impact from these reflections on the transmitted signal.

Note:

- Due to complications with part security we switched antennas.
- To be sure that the supported frequencies were what the data sheet specified, we tested the antennas S11 Return Loss.





Antenna Design Matrix

Antennas	Cost	Directionality (we want omnidirectional)	Gain at 600M in dBi (min average gain, max peak gain)	Gain at 850MHz in dBi (min average gain, max peak gain)	Total Score
TG.35.8113 - Taoglas Apex II Wideband 5G/4G Dipole Terminal Antenna	\$17.31 3	10	(-3.5,-2) 5	(-1.2,+1.6) 7	35
TG.45.8113- Taoglas Apex III Ultra-Wideband 5G/4G Dipole Terminal Antenna	\$16.46 4	10	(-4.2,0) 6	(-2,+2.7) 8	38 SOLD OUT
TG.46.8113 - Taoglas Apex IV Wideband 5G/4G Dipole Terminal Antenna	NA 0	10	(-3,-0.7) 6	(-3.5,+0.5) 6	32
TG.55.8113W - Taoglas 5G/4G Terminal Mount Monopole Antenna	\$5.77 8	(Need ideal ground plane) 3	(-4,+1) 6	(-4.5,+1) 6	31 43

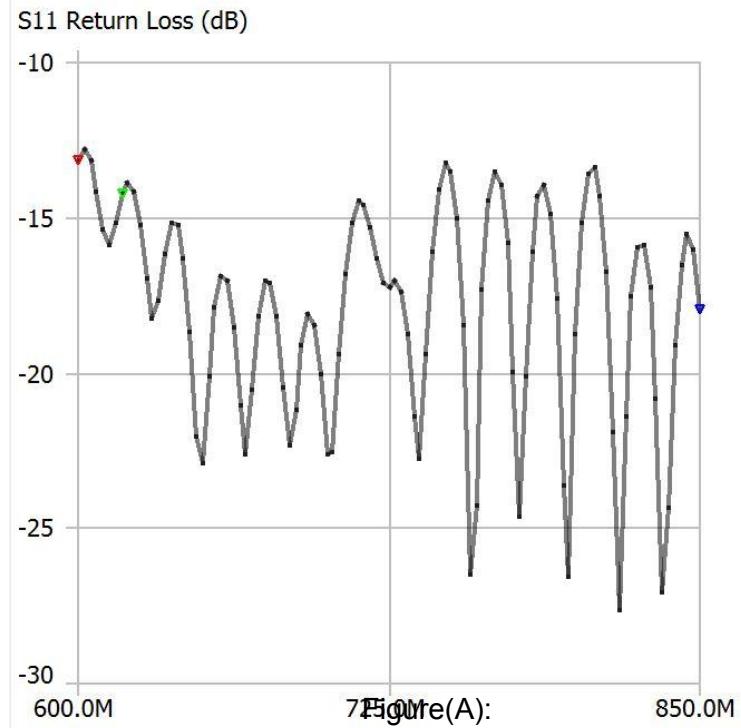
Test 3 - Results

Results:

- Figure(A) shows the return loss for the TG.35.8113 with a lossy coax cable.
 - The graph shows a large amount of oscillations making the data useless.

The Problem:

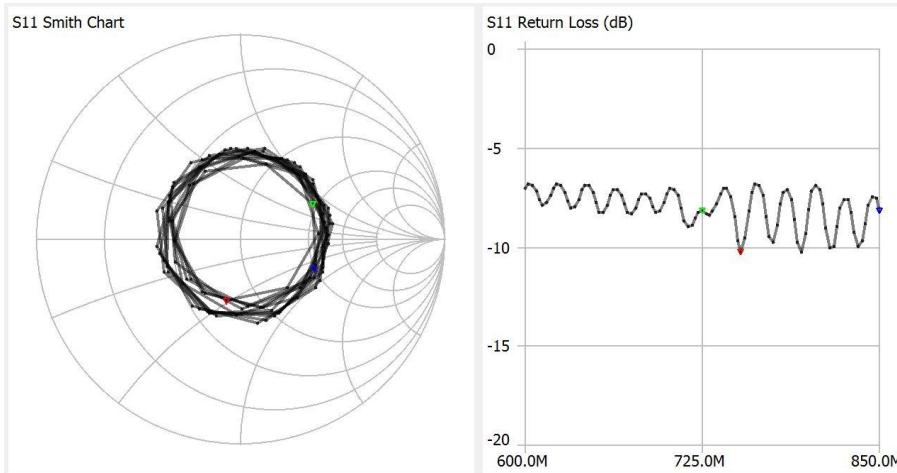
- This is not representative of the quality of the antenna, but rather the coax cable.



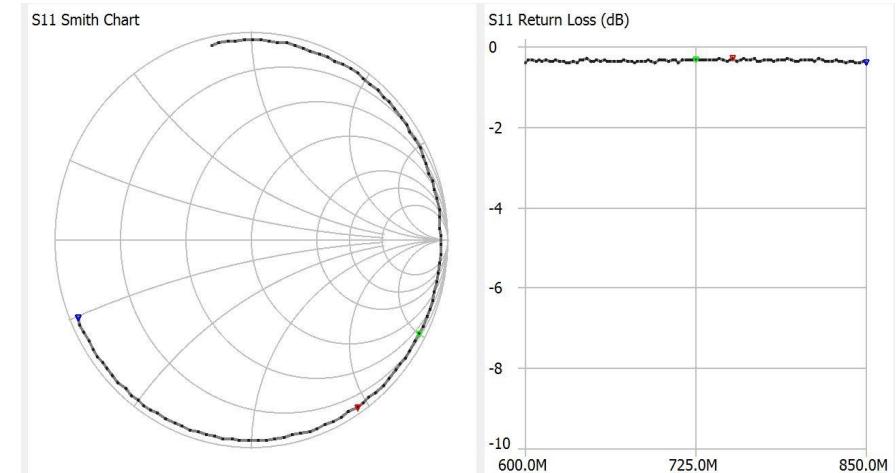
Test 3 - Results

The Problem Continued:

- The problem was pinpointed to be the coax cables the group ordered.
- It can be seen by comparing Figure(B) and Figure(C) where Figure(B) shows large amounts of oscillations and Figure(C) does not since it serves a high quality control in the experiment .



Figure(B):



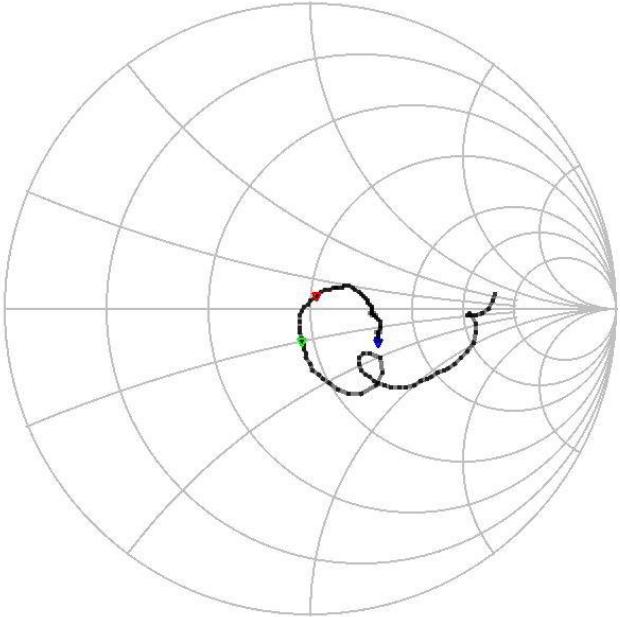
Figure(C):

Test 3 - Results

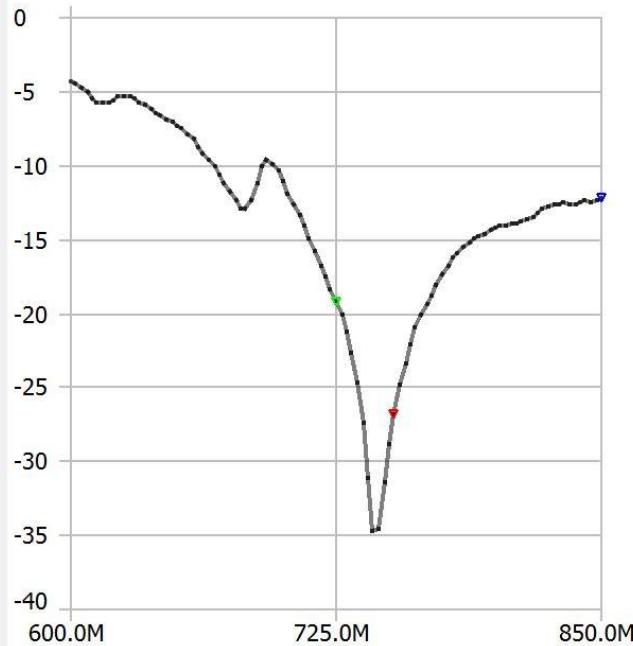
Final Results:

- Figure(D) shows the return loss for the TG.35.8113 with no coax cable.
- It can then be deduced by the graph that the ideal frequency to use is approximately around 750MHz

S11 Smith Chart



S11 Return Loss (dB)



Figure(D):