

# CAN WE RELY ON CITIZEN WEATHER MEASUREMENTS FOR OPERATIONAL USE?

Supervisors: Noel Fitzpatrick, Brandon Creagh

---

Matt Nagle

21st of September, 2021

# STRUCTURE OF PRESENTATION

1. What are Citizen Weather Measurements?
2. Can we trust them?
3. Quality control checks.
4. Project Summary.
5. Conclusions

# CITIZEN WEATHER MEASUREMENTS

The Weather Observations Website (WOW) is a **global network of observations from the public.**

# CITIZEN WEATHER MEASUREMENTS

The Weather Observations Website (WOW) is a **global network of observations from the public.**

Enthusiasts can purchase a **personal weather station** and upload measurements for and .

# CITIZEN WEATHER MEASUREMENTS

The Weather Observations Website (WOW) is a **global network of observations from the public.**

Enthusiasts can purchase a **personal weather station** and upload measurements for \_\_\_\_\_ and \_\_\_\_\_.

Since 2011, over 10,000 weather stations have reported over **1 billion weather observations worldwide.**

# OPERATIONAL USE CASE

- 91 Official Met Éireann stations (May 2021).
- 85 WOW stations (May 2021).

Potential Use Cases:

# OPERATIONAL USE CASE

- 91 Official Met Éireann stations (May 2021).
- 85 WOW stations (May 2021).

So, if we could trust the citizen weather measurements we could get a **93.4% increase in the unique measurements.**

Potential Use Cases:

# OPERATIONAL USE CASE

- 91 Official Met Éireann stations (May 2021).
- 85 WOW stations (May 2021).

So, if we could trust the citizen weather measurements we could get a **93.4% increase in the unique measurements.**

Potential Use Cases:

- Numerical Weather Prediction (**NWP**) Model verification.



# OPERATIONAL USE CASE

- 91 Official Met Éireann stations (May 2021).
- 85 WOW stations (May 2021).

So, if we could trust the citizen weather measurements we could get a **93.4% increase in the unique measurements.**

## Potential Use Cases:

- Numerical Weather Prediction (**NWP**) **Model verification.**
- NWP Data Assimilation - additional observation points for **establishing model initial conditions.**

# OPERATIONAL USE CASE

- 91 Official Met Éireann stations (May 2021).
- 85 WOW stations (May 2021).

So, if we could trust the citizen weather measurements we could get a **93.4% increase in the unique measurements.**

## Potential Use Cases:

- Numerical Weather Prediction (**NWP**) **Model verification.**
- NWP Data Assimilation - additional observation points for **establishing model initial conditions.**
- More dense network for **tracking severe weather.**

# OPERATIONAL USE CASE

---

# OPERATIONAL USE CASE

---

# CAN WE TRUST THE MEASUREMENTS?

---

# CAN WE TRUST THE MEASUREMENTS?

---

# CAN WE TRUST THE MEASUREMENTS?

- All WOW stations are **privately owned**.
- Met Éireann has **no control over the set up/calibration** of the devices.
- This can lead to significant measurement errors.
- Sometimes errors are blatant, other times it is not so clear.

# CAN WE TRUST THE MEASUREMENTS?

- All WOW stations are **privately owned**.
- Met Éireann has **no control over the set up/calibration** of the devices.
- This can lead to significant measurement errors.
- Sometimes errors are blatant, other times it is not so clear.

So, before operational use it is necessary to perform **comprehensive quality control checks** to remove errors.



# CAN WE TRUST THE MEASUREMENTS?

- All WOW stations are **privately owned**.
- Met Éireann has **no control over the set up/calibration** of the devices.
- This can lead to significant measurement errors.
- Sometimes errors are blatant, other times it is not so clear.

So, before operational use it is necessary to perform **comprehensive quality control checks** to remove errors.

There are two main types of checks:

1. Range Checks
2. Spatial Consistency Checks

## CHECK TYPE 1: RANGE CHECK

---

This involves checking if the **observation is within a chosen allowable range.**

# CHECK TYPE 1: RANGE CHECK

This involves checking if the **observation is within a chosen allowable range**.

Example 1: Feasibility Check

Max Temp (post 20th Century): 32.5 ;

Min Temp (post 20th Century): 18.8 ;

# CHECK TYPE 1: RANGE CHECK

This involves checking if the **observation is within a chosen allowable range**.

Example 1: Feasibility Check

Max Temp (post 20th Century): 32.5 ;

Min Temp (post 20th Century): 18.8 ;

So, if a measurement occurs outside of [ 20;35] it should be flagged as suspicious.

# CHECK TYPE 1: RANGE CHECK

This involves checking if the **observation is within a chosen allowable range**.

Example 1: Feasibility Check

Max Temp (post 20th Century): 32.5 ;

Min Temp (post 20th Century): 18.8 ;

So, if a measurement occurs outside of [ 20;35] it should be flagged as suspicious.

Example 2: We can use **historical climate values** to perform a range check for a **specific location** and **time**.

## CHECK TYPE 2: SPATIAL CONSISTENCY CHECK

---

The spatial consistency checks all rely on the same basic assumption:

Observations are spatially correlated on scales that can be observed by the network.

## CHECK TYPE 2: SPATIAL CONSISTENCY CHECK

---

The spatial consistency checks all rely on the same basic assumption:

Observations are spatially correlated on scales that can be observed by the network.

i.e. Stations that are close to each other should have similar observations.

# HOW DO WE DO THIS AUTOMATICALLY?

Humans are good at doing this .

We must explicitly tell a computer to:



# HOW DO WE DO THIS AUTOMATICALLY?

Humans are good at doing this .

We must explicitly tell a computer to:

1. Define a **neighbourhood** around the station being checked.

# HOW DO WE DO THIS AUTOMATICALLY?

Humans are good at doing this .

We must explicitly tell a computer to:

1. Define a **neighbourhood** around the station being checked.
2. Excluding the station being checked, calculate the **mean neighbourhood value** and a measure of the **spread of values**.

# HOW DO WE DO THIS AUTOMATICALLY?

Humans are good at doing this .

We must explicitly tell a computer to:

1. Define a **neighbourhood** around the station being checked.
2. Excluding the station being checked, calculate the **mean neighbourhood value** and a measure of the **spread of values**.
3. **Test how well the station being checked agrees with the neighbours.**

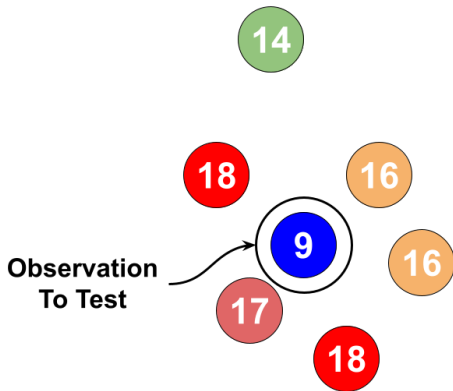
# HOW DO WE DO THIS AUTOMATICALLY?

Humans are good at doing this .

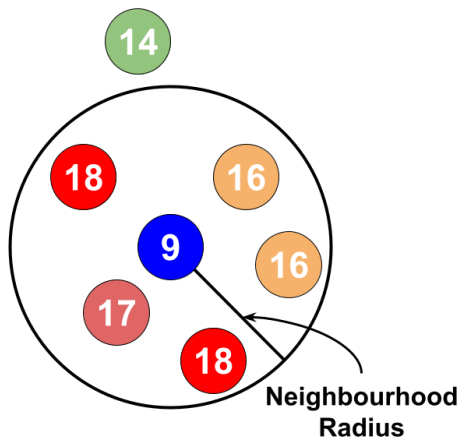
We must explicitly tell a computer to:

1. Define a **neighbourhood** around the station being checked.
2. Excluding the station being checked, calculate the **mean neighbourhood value** and a measure of the **spread of values**.
3. **Test how well the station being checked agrees with the neighbours.**
4. Based on a threshold **decide if the observation should be flagged.**

# SPATIAL CONSISTENCY CHECK - AUTOMATIC

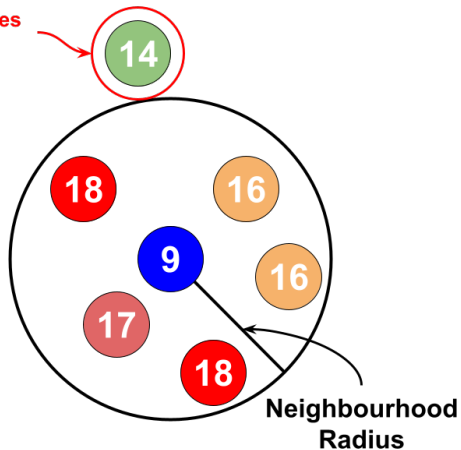


# SPATIAL CONSISTENCY CHECK - AUTOMATIC

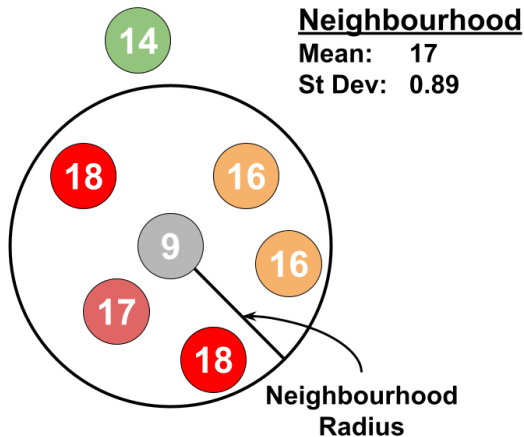


# SPATIAL CONSISTENCY CHECK - AUTOMATIC

Too Far Away - Does not affect check



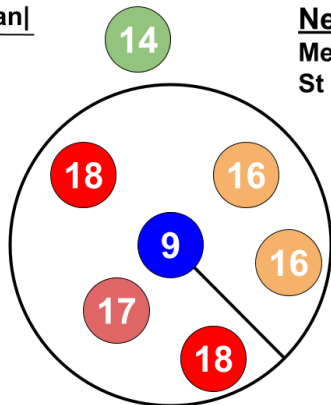
# SPATIAL CONSISTENCY CHECK - AUTOMATIC





# SPATIAL CONSISTENCY CHECK - AUTOMATIC

$$Z = \frac{|\text{Obs} - \text{Mean}|}{\text{St Dev}}$$



**Neighbourhood**

Mean: 17

St Dev: 0.89

# SPATIAL CONSISTENCY CHECK - AUTOMATIC

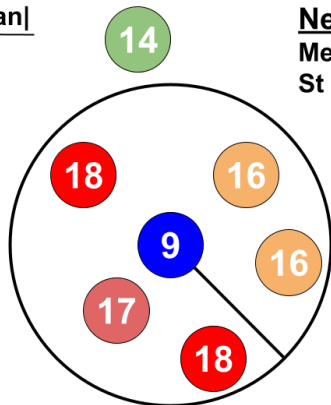
$$Z = \frac{|\text{Obs} - \text{Mean}|}{\text{St Dev}}$$

$$Z = \frac{|9 - 17|}{0.89}$$

Neighbourhood

Mean: 17

St Dev: 0.89

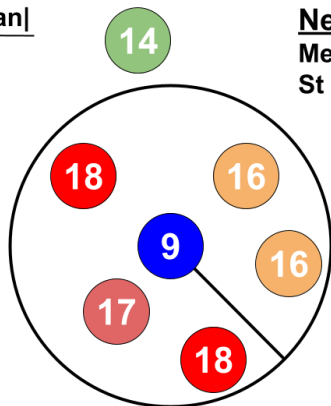


# SPATIAL CONSISTENCY CHECK - AUTOMATIC

$$Z = \frac{|\text{Obs} - \text{Mean}|}{\text{St Dev}}$$

$$Z = \frac{|9 - 17|}{0.89}$$

$$Z = 8.99$$



Neighbourhood

Mean: 17

St Dev: 0.89

# SPATIAL CONSISTENCY CHECK - AUTOMATIC

$$Z = \frac{|\text{Obs} - \text{Mean}|}{\text{St Dev}}$$

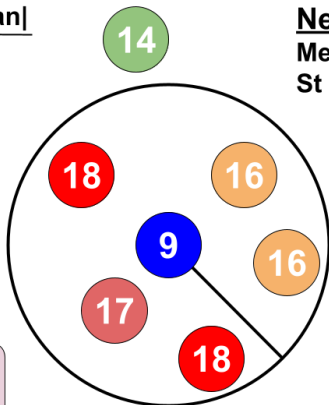
$$Z = \frac{|9 - 17|}{0.89}$$

$$Z = 8.99$$

## Neighbourhood

Mean: 17

St Dev: 0.89



If  $z > \sigma_{\text{thresh}}$   
 $\Rightarrow$  **Flag**

# COMPLICATION WITH IRISH DATA

Spatial Consistency Checks rely on on **data with large amounts of redundancy**.

We **want stations to be close to one another** - high density.

# COMPLICATION WITH IRISH DATA

Spatial Consistency Checks rely on on **data with large amounts of redundancy**.

We **want stations to be close to one another** - **high density**.

**Ireland has a low density of stations**. This makes spatial checks difficult to perform for the majority of stations.

# COMPLICATION WITH IRISH DATA

Spatial Consistency Checks rely on on **data with large amounts of redundancy**.

We **want stations to be close to one another** - **high density**.

**Ireland has a low density of stations**. This makes spatial checks difficult to perform for the majority of stations.

# MET NORWAY: TITAN SOFTWARE

- Luckily, we do not have to implement these checks from scratch.
- Met Norway have created and shared TITAN, software that contains functions to perform various Range & Spatial Consistency Checks.



# MET NORWAY: TITAN SOFTWARE

- Luckily, we do not have to implement these checks from scratch.
- Met Norway have created and shared **TITAN, software that contains functions to perform various Range & Spatial Consistency Checks.**
- Written in C++ but can be imported as a **Python or R package.**

# MET NORWAY: TITAN SOFTWARE

- Luckily, we do not have to implement these checks from scratch.
- Met Norway have created and shared **TITAN, software that contains functions to perform various Range & Spatial Consistency Checks.**
- Written in C++ but can be imported as a **Python or R package.**
- **Note:** Software is currently in an alpha stage.

## Testing:

1. Background research & Troubleshooting TITAN installation.
2. Extensive QC testing on **May 2021 Official & WOW station data.**

## Outputs:

## Testing:

1. Background research & Troubleshooting TITAN installation.
2. Extensive QC testing on **May 2021 Official & WOW station data**.

## Outputs:

1. **Documentation** for the TITAN software.

## Testing:

1. Background research & Troubleshooting TITAN installation.
2. Extensive QC testing on **May 2021 Official & WOW station data**.

## Outputs:

1. **Documentation** for the TITAN software.
2. Jupyter Notebooks **Examples for each TITAN check** were created.

## Testing:

1. Background research & Troubleshooting TITAN installation.
2. Extensive QC testing on **May 2021 Official & WOW station data**.

## Outputs:

1. **Documentation** for the TITAN software.
2. Jupyter Notebooks **Examples for each TITAN check** were created.
3. **Improvements to the existing TITAN range checks** to make them more applicable to Irish Data.

## Testing:

1. Background research & Troubleshooting TITAN installation.
2. Extensive QC testing on **May 2021 Official & WOW station data**.

## Outputs:

1. **Documentation** for the TITAN software.
2. Jupyter Notebooks **Examples for each TITAN check** were created.
3. **Improvements to the existing TITAN range checks** to make them more applicable to Irish Data.
4. Custom scripts to **apply the chosen checks sequentially and visualise the results** for easier interpretation.

# IMPROVEMENTS - CLIMATOLOGY RANGE CHECK

TITAN calculates a **guess for the mean temp based on Latitude.**



# IMPROVEMENTS - CLIMATOLOGY RANGE CHECK

TITAN calculates a **guess for the mean temp based on Latitude.**

We can use the **30 year normal climate values from observations.**

# CONCLUSIONS/RECOMMENDATIONS

- **WOW data is rich** and could be very useful for Met Éireann.

# CONCLUSIONS/RECOMMENDATIONS

- **WOW data is rich** and could be very useful for Met Éireann.
- **TITAN is powerful software written to perform Quality Control checks** on citizen weather measurements.

# CONCLUSIONS/RECOMMENDATIONS

- **WOW data is rich** and could be very useful for Met Éireann.
- **TITAN is powerful software written to perform Quality Control checks** on citizen weather measurements.
- After performing the improved TITAN range checks, **the WOW data can be used as a fairly reliable reference data set.**

# CONCLUSIONS/RECOMMENDATIONS

- **WOW data is rich** and could be very useful for Met Éireann.
- TITAN is powerful software written to perform Quality Control checks on citizen weather measurements.
- After performing the improved TITAN range checks, **the WOW data can be used as a fairly reliable reference data set.**
- In the future **if more stations are added we would be able to reliably perform the more complex checks** and have even more confidence in the resulting data set.

QUESTIONS?

## APPENDIX: IMPROVEMENTS - CLIMATOLOGY RANGE CHECK

TITAN calculates a **guess for the mean temp based on Latitude.**

## APPENDIX: IMPROVEMENTS - CLIMATOLOGY RANGE CHECK

TITAN calculates a **guess for the mean temp based on Latitude.**

We can use the **30 year normal climate values from observations.**



Base Air Temp Measurements.

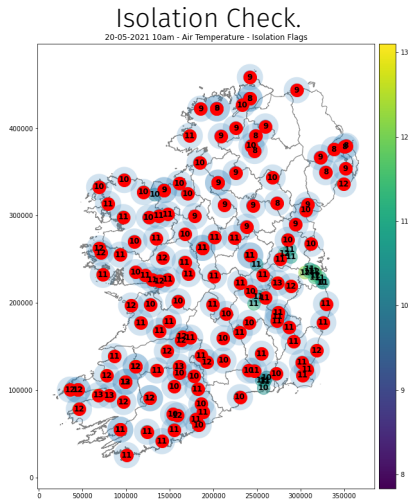
# FINAL PRODUCT - EXAMPLES

Base Air Temp Measurements.

Custom Climate Range Check.

# FINAL PRODUCT - EXAMPLES

Base Air Temp Measurements.



# FINAL PRODUCT - EXAMPLES

Base Air Temp Measurements.

Spatial Consistency Test.

