STRTEAMGAGE PREDICTION IN THE NORTHEAST

BY SVEN McCALL JR.

ABSTRACT

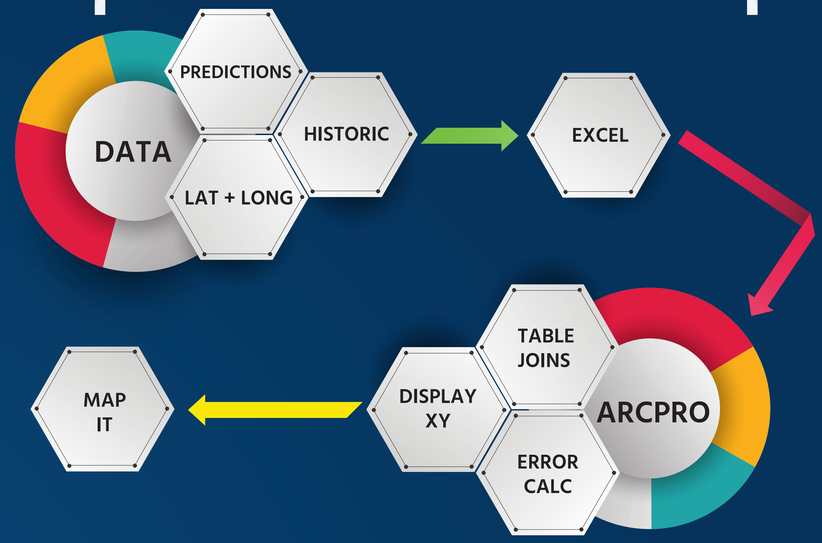
STREAMGAGES ARE TOOLS THAT MEASURES THE AMOUNT OF WATER MOVING THROUGH A RIVER OR A STREAM. THE USGS HAS OVER 1.5 MILLION OF THESE ACCROSS THE UNITED STATES. ONE GOAL IS TO BE ABLE TO USE DATA THAT THESE GAGES PROVIDE AND COME UP WITH EQUATIONS TO PREDICT THE PEAK RIVER FLOW A GAGE WILL ENCOUNTER IN THE FUTURE. THE USGS HAS AN ONLINE TOOL , CALLED STREAMSTATS, THAT USES REGRESSION EQUATIONS BUILT FROM PREVIOUS DATA TO MAKE THESE PREDICTIONS. ONE THING THAT STANDS OUT IS OFTEN TIMES THESE EQUATIONS WILL CHANGE WHEN STATE LINES ARE CROSSED. THE GOAL OF THIS PROJECT WAS TO SEE IF THE WAY THE EQUATIONS ARE BOUNDED (BY STATE LINES) SEEMS TO YIELD A PATTERN OF INCONSISTENCY. IN OTHER WORDS, CAN ALL THE PREDICTIONS BE LOOKED AT ACROSS MULTIPLE STATES AND VISUALIZED USING GIS TECHNOLOGY TO SHOW THAT THIS METHOD OF CHOOSING A REGRESSION EQUATION FOR A SITE IS SOMEWHAT BACKWARDS. THE IDEA IS THAT IF NATURE DOES NOT PAY ATTENTION TO STATE BOUNDARIES, WHY IS THAT WE ARE USING THESE BOUNDARIES TO PREDICT NATURAL EVENTS. WE PEER INTO DATA AND PREDICTION FOR SOMETHING THAT IS EXTREAMLY VARIABLE AND THUS NEEDS TO BE LOOKED AT CAREFULLY. THE RESULTS BELOW ARE FROM A RELATIVELY SMALL SET OF A MUCH LARGER NETWORK, BUT IF THE GROUND WORK CAN BE LAYED THEN MAYBE MORE AND MORE OF THESE STUDIES WILL YIELD INFORMATION THAT WILL CHANGE THE WAY WE THINK ABOUT STATE BOUNDARIES AND NATURAL EVENTS.

DATA



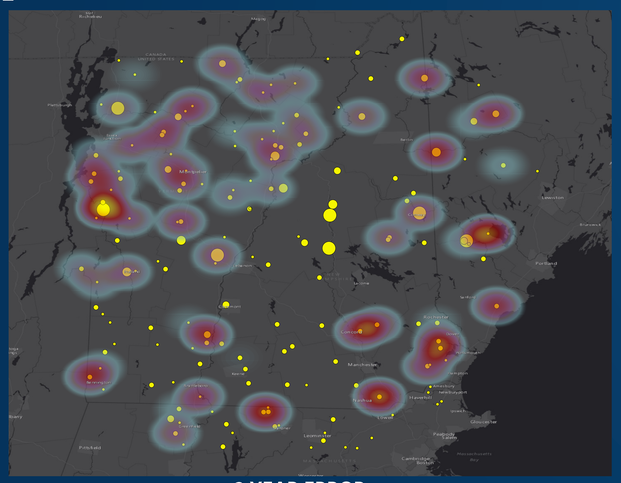
THE DATA IS THE KEY TO THIS ENTIRE ENDEVOUR. THE DATA IS EASY TO CALCULATE, PLENTIFUL AND HARD TO DRAW CONCLUSIONS FROM. AS THE WORK FLOW POINTS OUT, WE START WITH ALL THE SITES IN AN AREA THAT HAVE PREDICTIONS. ONCE THOSE ARE COLLECTED, ALL HISTORIC DATA FROM EACH USGS SITE NEEDS TO BE GATHERED, ORGANIZED IN EXCEL AND THEN SET UP TO TO START CALCULATING ERROR RATES. THE CALCULATION IS SIMPLE. WE USE AN ERROR PERCENTAGE FORMULA AND SIMPLY MAKE EVERY PERCENTAGE POSITIVE TO FOCUS ONLY ON PURE ERROR AND AVOID THE COMPLEXITY OF + OR - . ERROR PERCENTAGE WAS CHOSEN FOR ITS ABILITY TO NOT TAKE DISCHARGE SIZES INTO ACCOUNT. WE TREAT EACH GAGE AS THOUGH IT WERE THE SAME SIZE. HOWEVER, WE DO COME BACK TO DISCHARGE WITH DIFFERENT SYMBOLOGY IN THE FINAL MAPS.

WORK FLOW

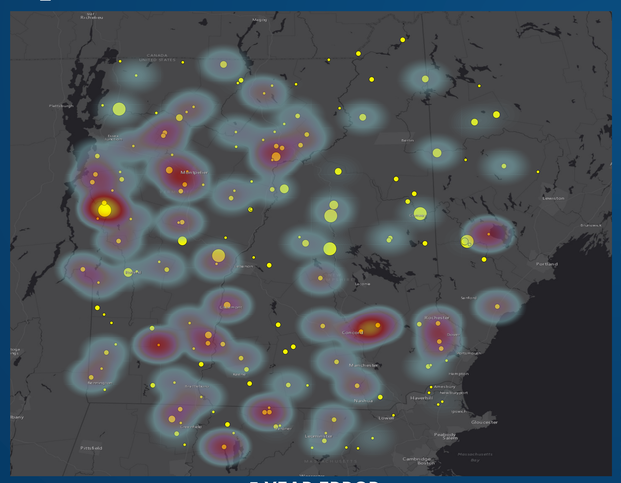


FINDINGS

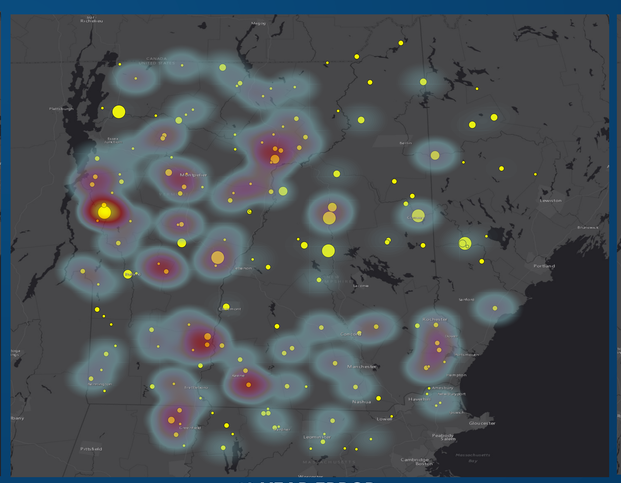
THE BEST POSSIBLE WAY TO REPRESENT THESE RESULTS WAS TO DISPLAY THEM IN A HEAT MAP SYMBOLOPGY. USING A HEAT MAP FOCUSED ON A SPECIFIC FIELD ALLOWS US TO WATCH TRENDS CHANGE OVER TIME. THESE MAPS ARE INTUITIVE AND READ LIKE YOU WOULD THINK. ZONES THAT ARE 'HOTTER' HERE REPRESENT DENSE AREAS THAT HAVE HIGH ERROR RATES. AS THE ERROR YEARS INCREASE, WE CAN START TO SEE WHERE CONSISTENCY HAS ARISEN. AREAS THAT START HOT IN THE 2 YEAR ERRORS AND CONTINUE THRU THE 25 YEAR ERRORS ARE PLACES THAT SEEM TO KEEP HAVING HIGH ERROR RATES, WHETHER THAT BE HIGH OR LOW. THE SECOND PIECE OF INFORMATION TO NOTICE ARE THE ACTUAL STREAMGAGE LOCATIONS AND THESE ARE LARGER IF THEY DEAL WITH MORE DISCHARGE OF WATER. THE BIG THING WE LOOK FOR IN THESE MAPS ARE 'HOT' AREAS WITH LARGE STREAMGAGE POINTS. THIS WOULD INDICATE ERRORS WHERE WE CERTAINLY DON'T WANT THEM. AS ALWAYS, LOOK AT THE MAPS KNOWING THAT SOME SPOTS DO NOT HAVE DATA WHICH WILL SHOW UP AS A YELLOW DOT WITH NOT SHADE OF BLUE AROUND IT AT ALL. THESE LOCATIONS DID NOT HAVE THE HISTORIC DATA TO COMPARE.



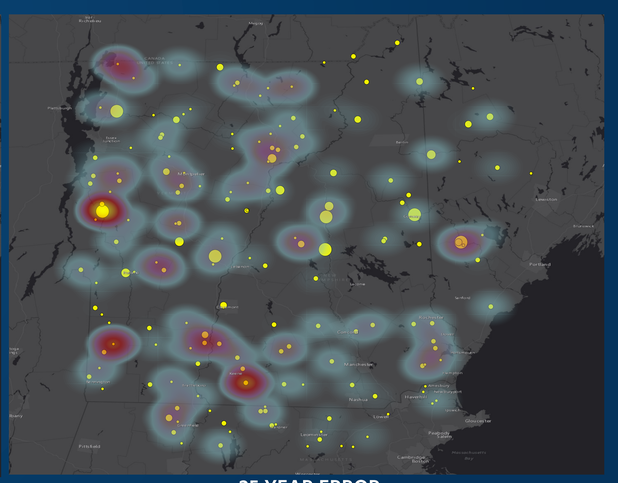
2 YEAR ERROR



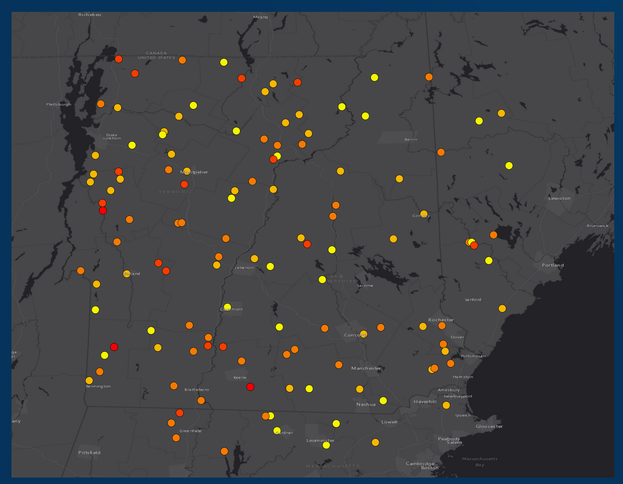
5 YEAR ERROR



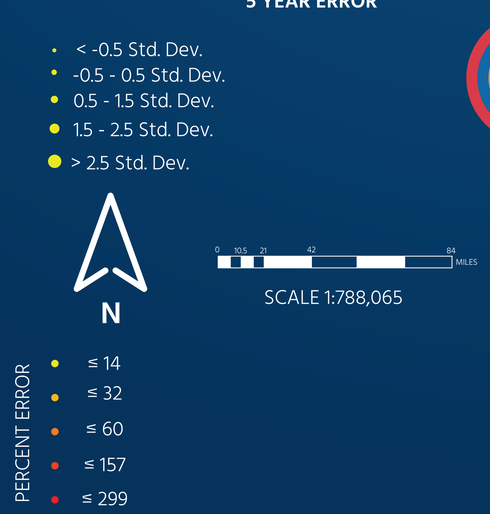
10 YEAR ERROR



25 YEAR ERROR



25 YEAR ERROR COLOR ONLY



SUMMARY

WHEN WE TAKE A GOOD LONG LOOK AT THE DATA THERE ARE SOME APPARENT AREAS WHERE WE SEE CONSISTENT HOT ZONES ON LARGER DISCHARGE ZONES WHICH WOULD BE DEVASTATING FOR FLOODING IN THE POSITIVE DIRECTION AND IT WOULD BE A WASTE OF RESOURCES IN THE NEGATIVE. THERE IS NO DOUBT THAT WHAT WE ARE SEEING HERE IS A SMALL ZONE OF THE MUCH LARGER PICTURE BUT WE REALLY SEE NO APPARENT TRENDS OF ZONING ISSUES. THESE PATTERNS WOULD PRESENT AS VERY DRASTIC CHANGES IN ERRORS NEAR STATE LINES. WE SEE SIMILAR PATTERNS CLUMPED NEAR STATE LINES WHICH SUGGESTS THAT THEY SEE SIMILAR AMOUNTS OF WATER AND THAT THE EQUATIONS ARE DOING WHAT THEY ARE SUPPOSE TO. HOWEVER, THE DATA DOES SHOW THAT THERE IS STILL A LONG WAY TO GO WITH STREAMGAGE PREDICTION BUT TO STEP IN THE RIGHT DIRECTION, STUDIES LIKE THESE NEED TO BE DONE. THE NEXT STEP IN THE PROCESS WOULD BE TO START TO PUSH GAGES INTO ZONES BASED ON THEIR SURROUNDING GEOGRAPHY AND IF WE SEE A STATISTICALLY SIGNIFICANT RESULT THAT THERE IS A PATTERN OF SHARED ERROR PERCENTAGES, THEN THAT WOULD SUGGEST A NEW ZONING METHOD THAT WOULD BE FAR MORE ACCURATE. SINCE ONLY ONE VARIABLE GOES INTO WHICH EQUATION TO USE (THE STATE IN WHICH A GAGE LIES) , ADDING A SECOND LEVEL OF COMPLEXITITY BASED ON SURROUNDING GEOGRAPHY SHOULD MAKE THESE ERROR PERCENTAGES DROP SIGNIFICANTLY.

SOURCES

USGS science for a changing world

esri

The UNIVERSITY of VERMONT