Estimating Plant Available Water for General Crop Simulations in Almanac/APEX/EPIC/SWAT

K.D. Behrman, J.R. Willams, R.A.J. Taylor, J.R. Kiniry & M.L. Norfleet

USDA-ARS, USDA-NRCS & Texas A&M University
Blackland Research Center
Temple, Texas, U.S.A.

Plant Available Water

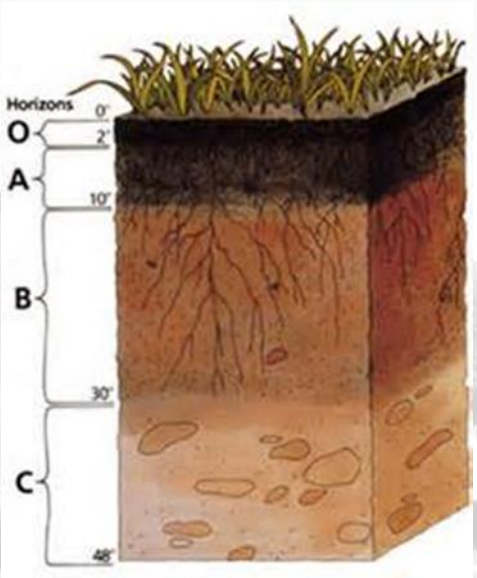
- A critical variable in the plant growth modules of Almanac, EPIC/APEX and SWAT is soil moisture.

- Plant Available Water is a fundamental driver of the plant sub models.

- PAW is computed from the difference between Field Capacity and Wilting Point (PAW = FC-WP).

Soil Data for Modeling

- All four models use soils data stored in the STATSGO and SSURGO databases and National Soil Characterization Laboratory database.
- Data are stored by horizon.
- FC & PW are horizon-specific fields in these databases.



There are some inconsistencies in the these databases:

a. The survey, analytical and reporting methods and standards have changed since the first soil data were collected in 1942, but the databases use the original results;

b. There are real differences between variables of similar soils depending on when and where they were collected.

Leading to problems …

• Output from simulations using different soils may be confounded with the changes in survey, analysis or recording methods that have evolved over 70 years.

• Plant (crop) growth is an important driver of soil evolution.

• Flaws in the plant growth projections can have profound effects on the conclusions drawn from model scenarios.

Solution Part 1

• USDA-NRCS developed a database of 39,366 records of soil field capacity (FC) and wilting point (WP) in relation to soil texture (%Clay & %Sand) and organic carbon (%OC).

• Compare current model soil conditions against the reference set, and

• Compute Plant Available Water (PAW = FC – WP) for use in plant growth models.

Solution Part 2

Attilla Nemes* developed a “brute force” Nearest Neighbor method for finding FC and WP from %Clay, %Sand, %OC using the NRCS database:

• Calculate the Euclidian distance between target and reference points;

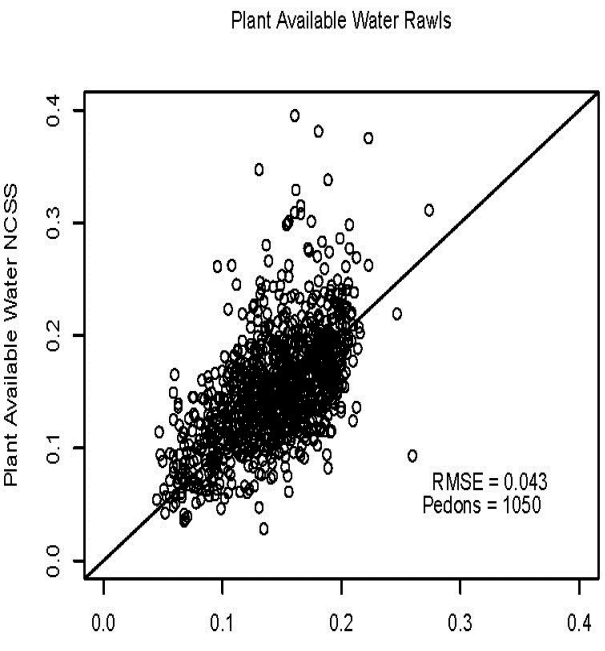
• Sort the distances in ascending order;

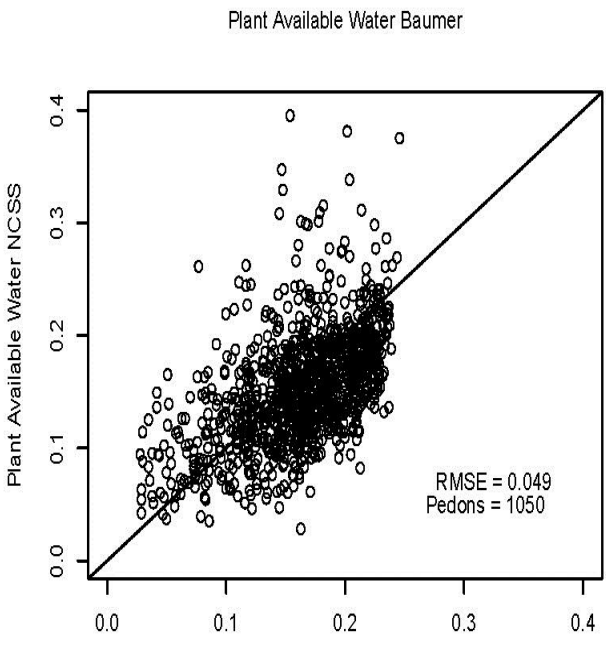
• Average the response variables (FC & WP) for the *N* nearest neighbors.

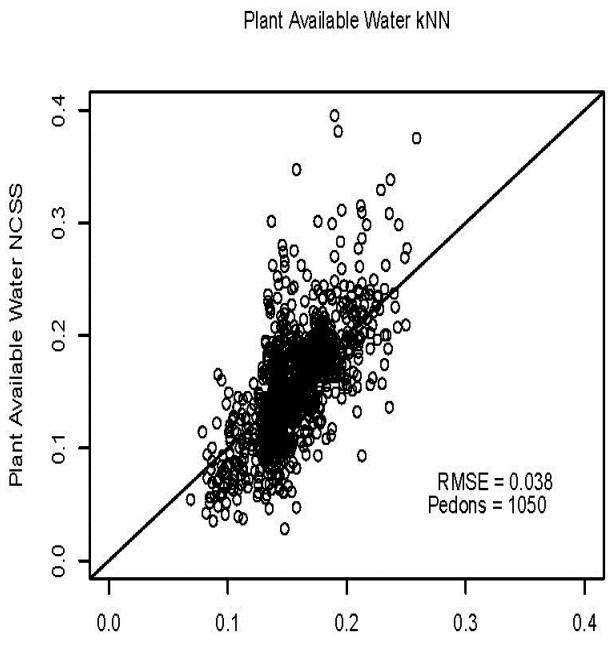
* Nemes et al. *Soil Sci. Soc. Am. J.* **70**:327–336 (2006); *Vadose Zone J.* **5**:1222–1235 (2006).

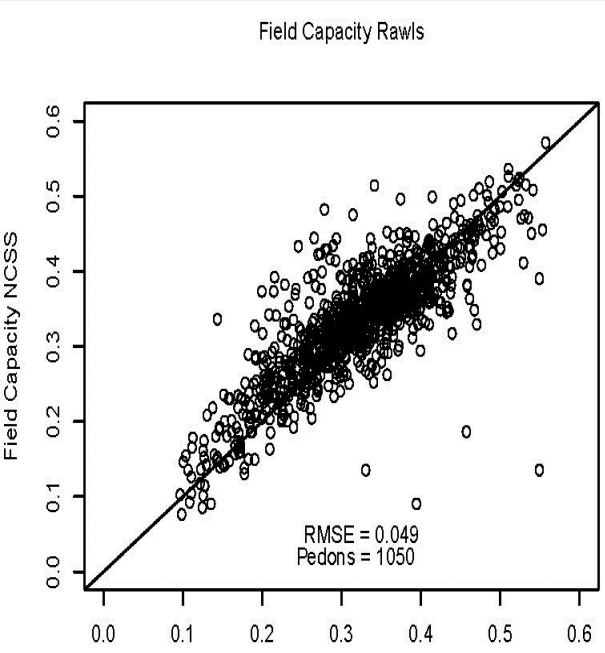
Nearest Neighbor vs. Computed PAW

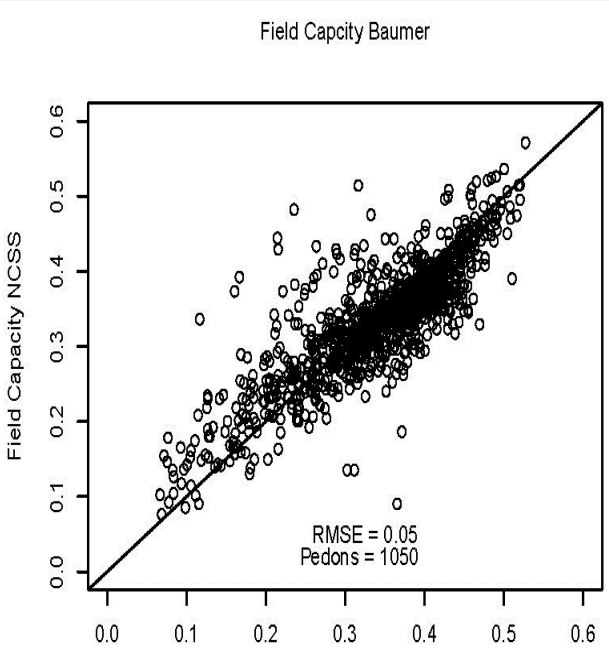
The Nearest Neighbor method gives better estimates than the computational process methods of Baumer et al. (1994) and Rawls et al. (1991), based on RMSE tested using 1054 different soils not in the database.

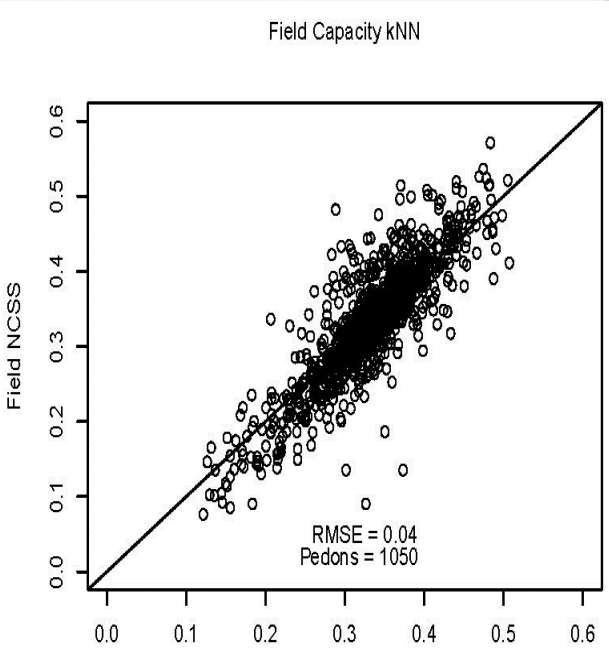


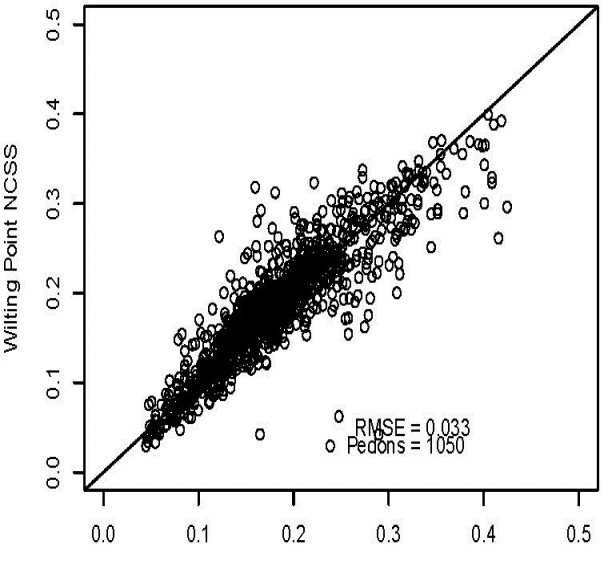


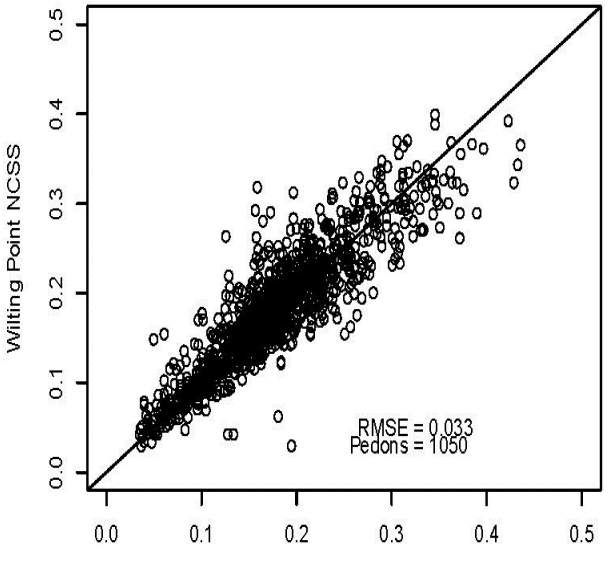


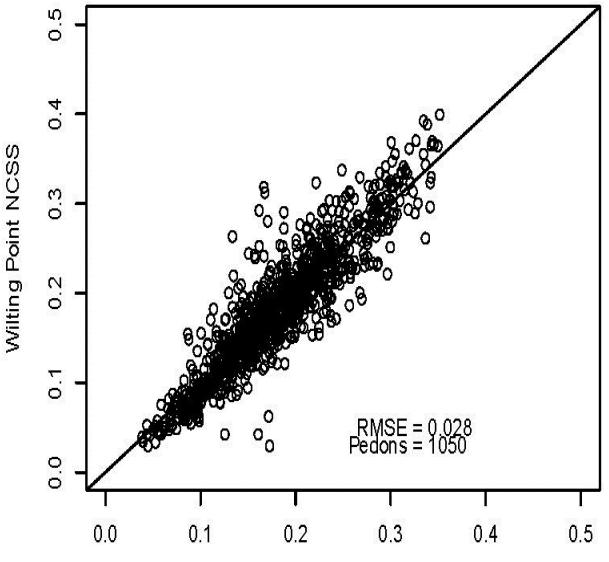

Field Capacity Rawls
Field Capacity NCSS
RMSE = 0.049
Pedons = 1050


Field Capcity Baumer
Field Capacity NCSS
RMSE = 0.05
Pedons = 1050


Field Capacity kNN
Field NCSS
RMSE = 0.04
Pedons = 1050


Wilting Point Rawls
Wilting Point NCSS
RMSE = 0.033
Pedons = 1050


Wilting Point Baumer
Wilting Point NCSS
RMSE = 0.033
Pedons = 1050


Wilting Point kNN
Wilting Point NCSS
RMSE = 0.028
Pedons = 1050

Unfortunately, the nearest neighbor method is very slow: it takes approximately 14X as long as the computational methods:

Benchmarks for 1054 40-yr Simulations

	Hrs	Mins	Secs	Secs/run	Runs/min
Baumer (1994)	0	14	42.68	0.413	145.3
Rawls (1991)	0	15	9.14	0.425	141.1
Nearest Neighbor	3	22	31.27	5.683	10.6

Rawls et al. (1991) *Advances in Soil Science* **16**:213-234.
Baumer et al. (1994) *MUUF v2.14 User's Manual.* USDA-NRCS.

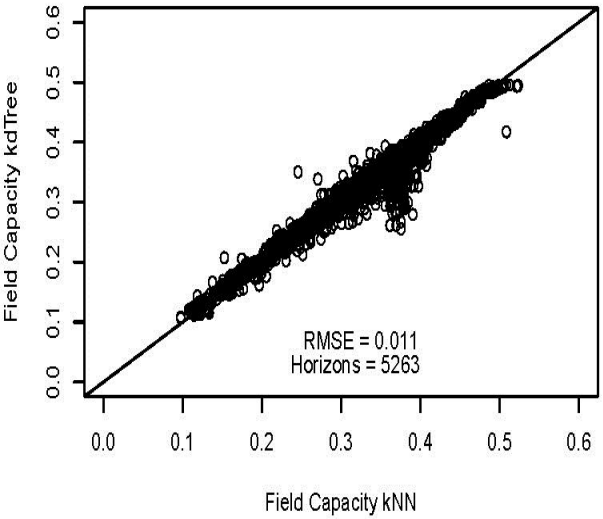
The K-D Tree Binary Search

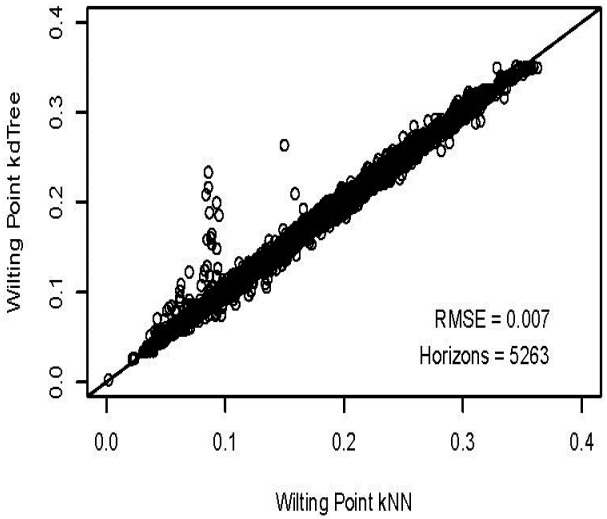
A solution is to continue to use the nearest neighbor method, but use a non-parametric technique called K-D Trees* to find the nearest neighbors in *K*-dimensions more efficiently.

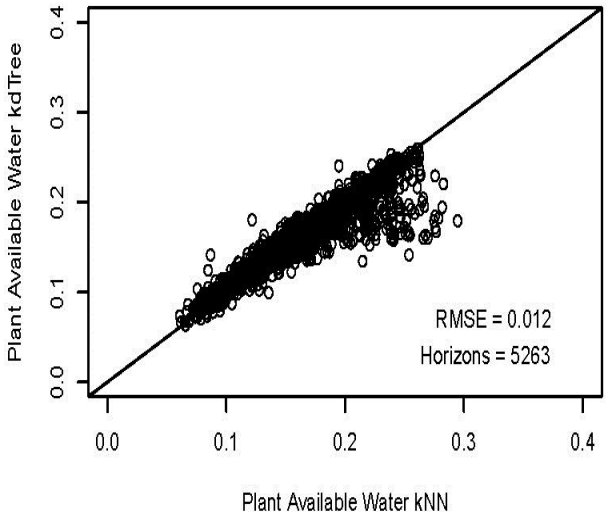
K-D Tree is a binary search method in $K < 20$ dimensions that operates in $K \log N$ time.

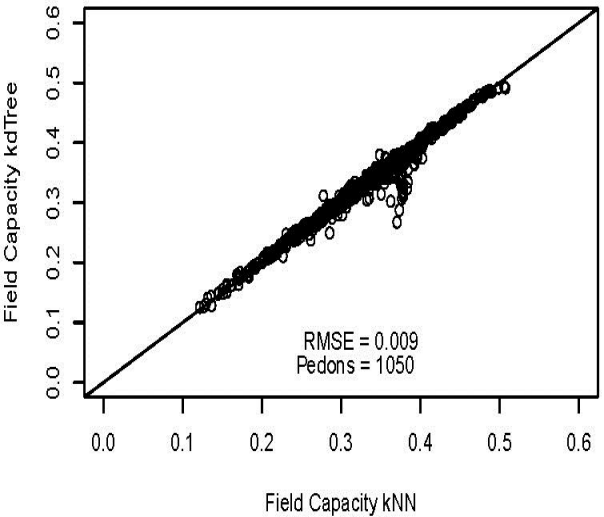
* Friedman et al. (1977) *ACM Trans. Math. Soft.* B:209-226.

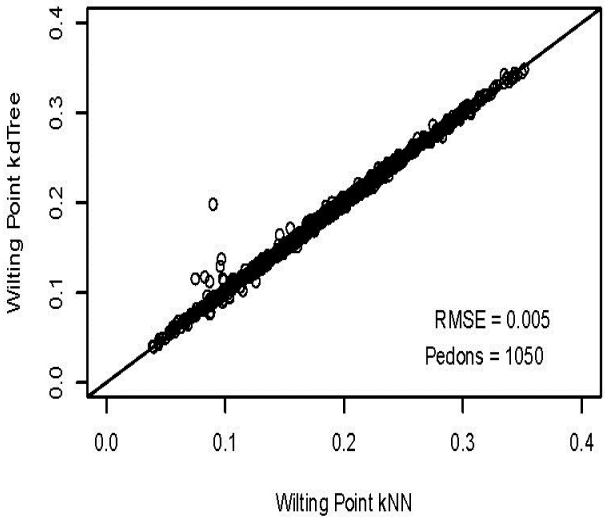
Comparing NN & K-D Tree:

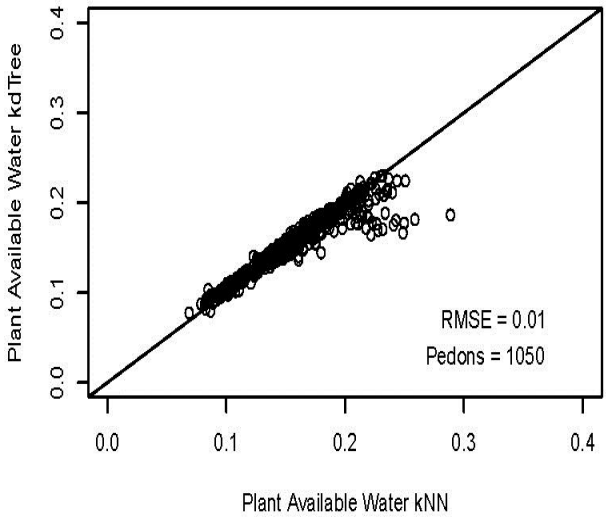












How the K-D Tree Compares …

Benchmark Time

	Hrs	Mins	Secs	Secs/run	Runs/min
Baumer (1994)	0	14	42.68	0.413	145.3
Rawls (1991)	0	15	9.14	0.425	141.1
Norfleet (unpubl.)	0	15	15.11	0.428	140.2
Nearest Neighbor	3	22	31.27	5.683	10.6
KD-Tree (Friedman 1977)*	0	21	10.29	0.594	101.0

•Time to build K-D tree: 0.156 sec

Friedman et al. (1977) *ACM Trans. Math. Soft.* B:209-226
Rawls et al. (1991) *Advances in Soil Science* **16**:213-234.
Baumer et al. (1994) *MUUF v2.14 User's Manual. USDA*-NRCS.

Conclusions

- K-D Tree method for looking up soil properties is effective and efficient.

- Lookup tables are a viable means of “filling in” missing values in databases.

- Are there other properties that could be looked up like this?