MICROWAVE REMOTE SENSING-BASED MACHINE LEARNING METHOD FOR IRRIGATION ESTIMATION IN FLORIDA

Laura Almendra-Martín[1], Jasmeet Judge[1], Alejandro Monsivais-Huertero[2], Pang-Wei Liu[3], George Worrall[4]

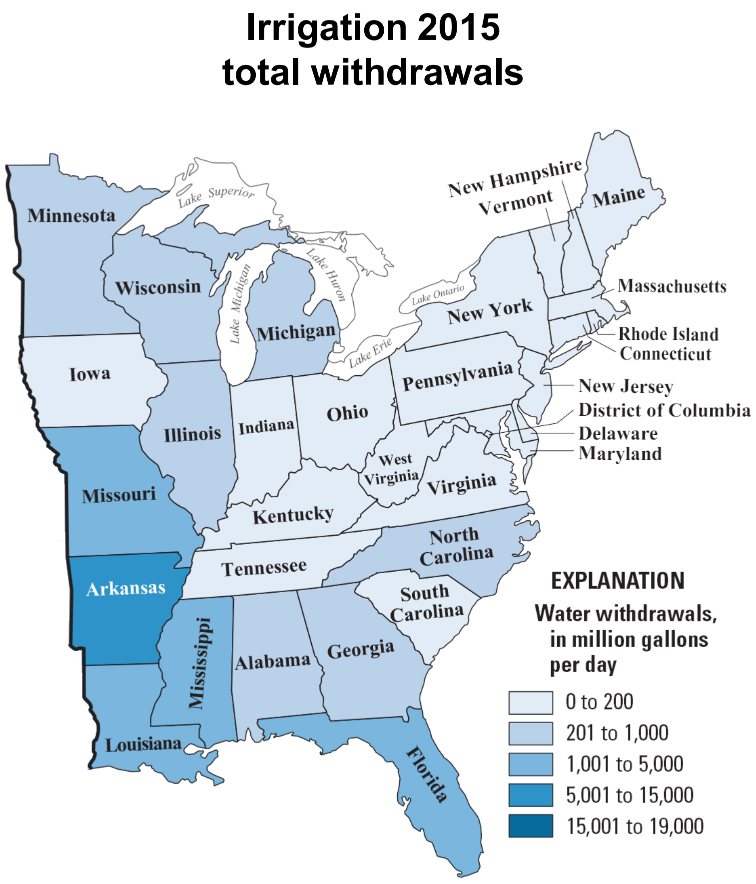
[1]Center for Remote Sensing, Agricultural and Biological Engineering Department, University of Florida
[2]Instituto Politécnico Nacional, Mexico City, Mexico
[3]NASA Goddard Space Flight Center, Maryland
[4]Deep Grain, St Louis, Missouri

FUNDINGS SUPPORT FROM USDA-CIG AWARD #P0206601

CONTENT

1. **Introduction and motivation**
    - Significance of irrigation estimation in Florida.
    - Microwave remote sensing and hydrology.
2. **Objectives**
3. **Methods**
    - Synthetic data generation.
    - Machine learning framework: training and validation.
4. **Results**
5. **Summary and future work**

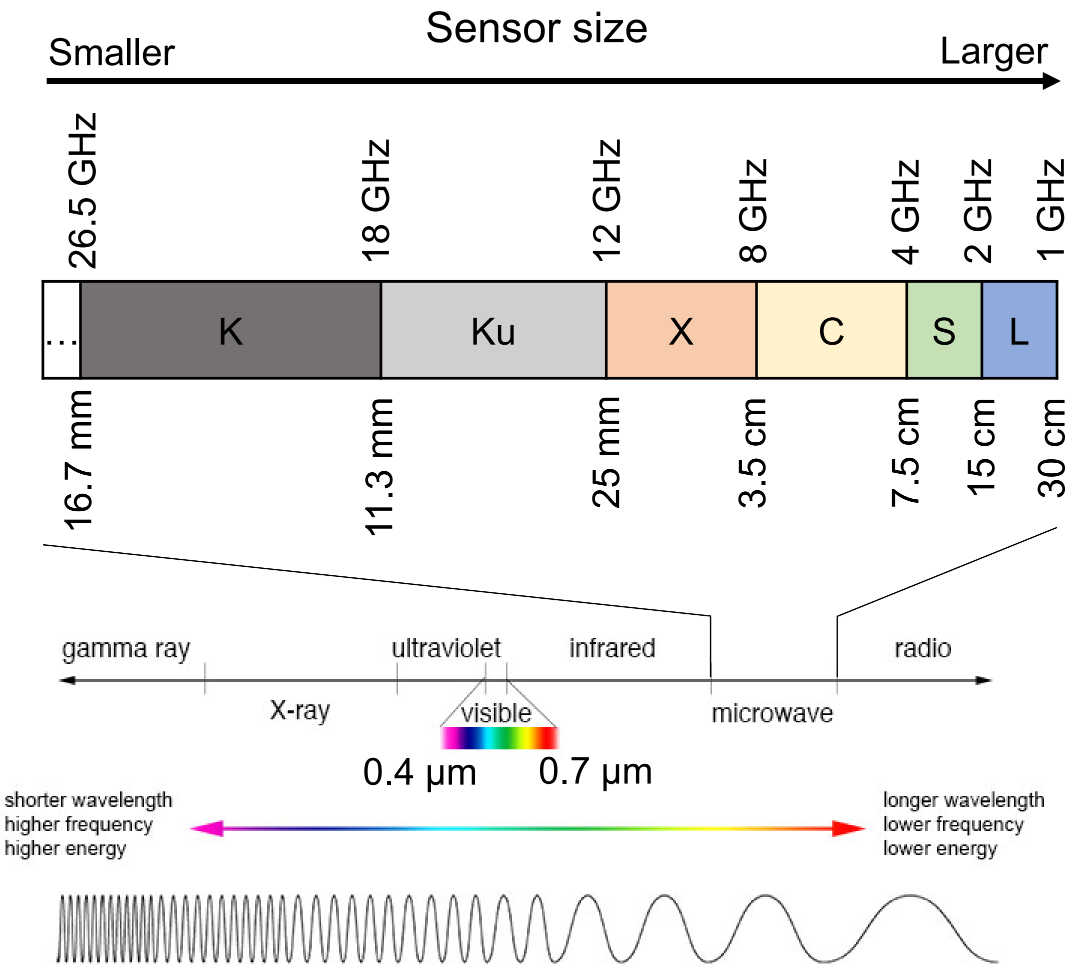
IRRIGATION IN FLORIDA



Dieter aet al., 2018. Estimated use of water in the United States in 2015: U.S. Geological Survey Circular

- Irrigation records typically come from surveys.
- Monitoring this activity at high spatio-temporal scales is challenging but can ensure an efficient water usage.
- Florida is one of the most heavily irrigated areas in eastern US.
- Soil moisture (SM) has shown to be a key variable for irrigation estimations.

MICROWAVE REMOTE SENSING



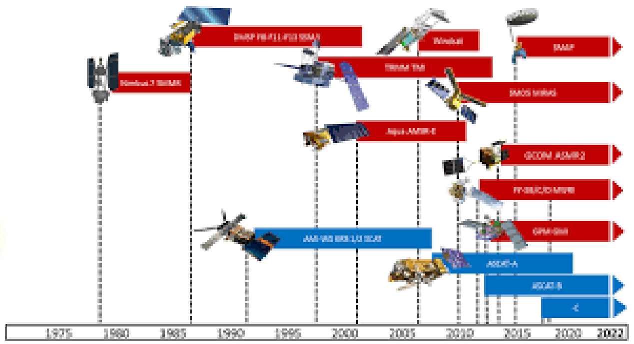
(Credit: NASA's Imagine the Universe)

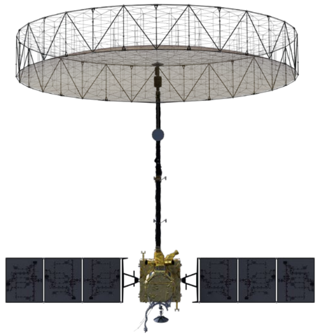
- Microwave penetrate through clouds, vegetation, and soil.
- Allows day-and-night observation.
- Interaction of microwaves with materials depends on their dielectric properties.
- SM can be accurately retrieved using microwave remote sensing sensors.

REMOTE SENSING SM PRODUCTS

- Early microwave-based SM products available globally every 2-3 days at 40-36 km spatial resolution.

- Downscaling techniques → up to 1 km SM product.
- Merging data from different satellites generate long-term global SM records.

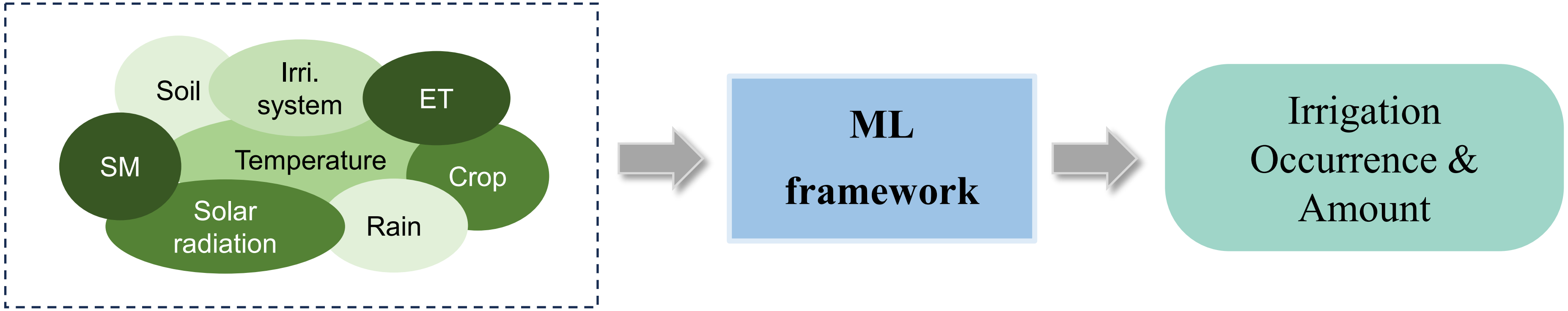




- Upcoming missions are expected to generate SM products at 100 to 200 m.

OBJECTIVES

Develop a machine learning (ML) framework to estimate irrigation occurrence and amount in Florida.

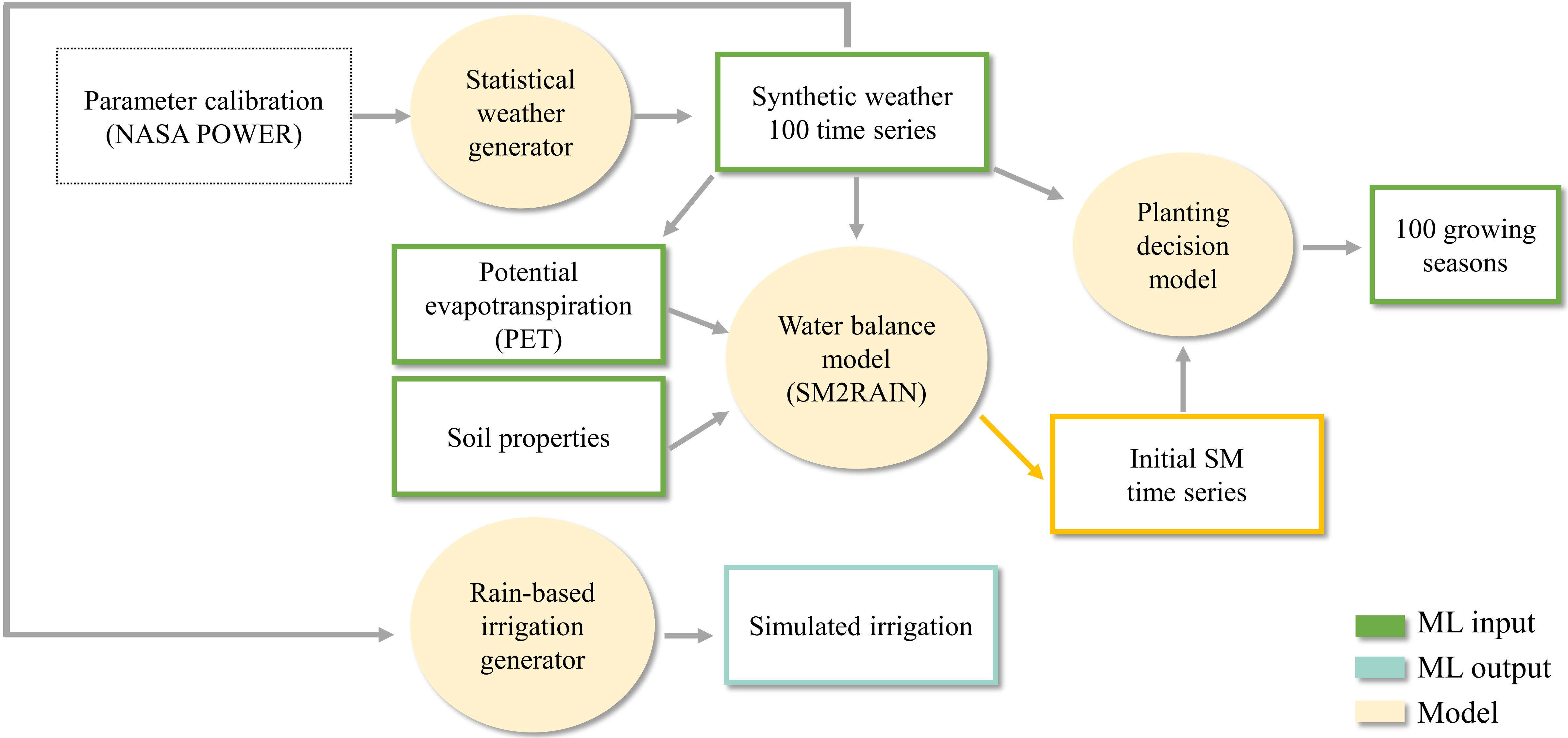


ML algorithms need data for training, but irrigation records are scarce.

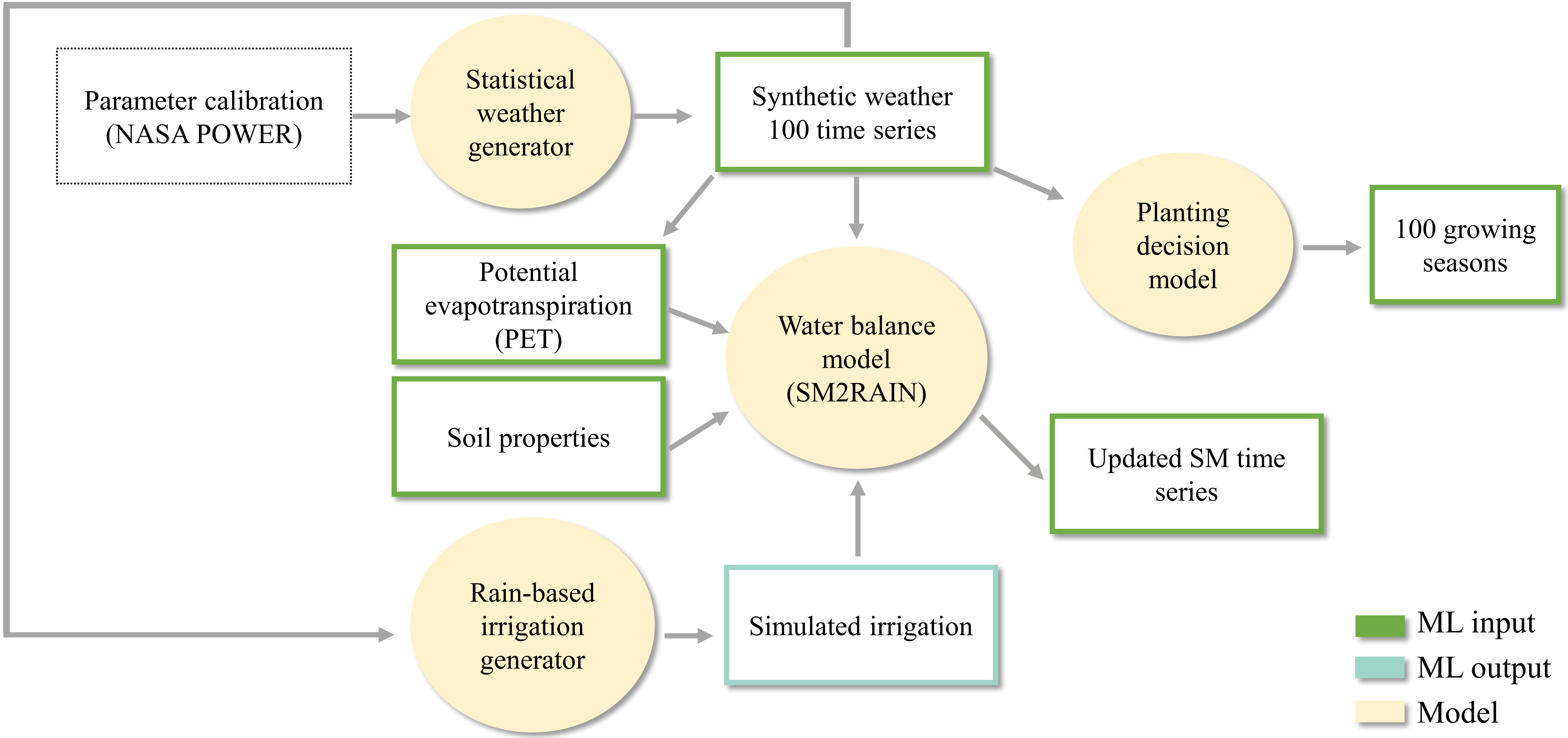
I. Generate synthetic data using physically-based models for training.

II. SVM-based ML for irrigation estimation for future high-resolution microwave-based SM products.

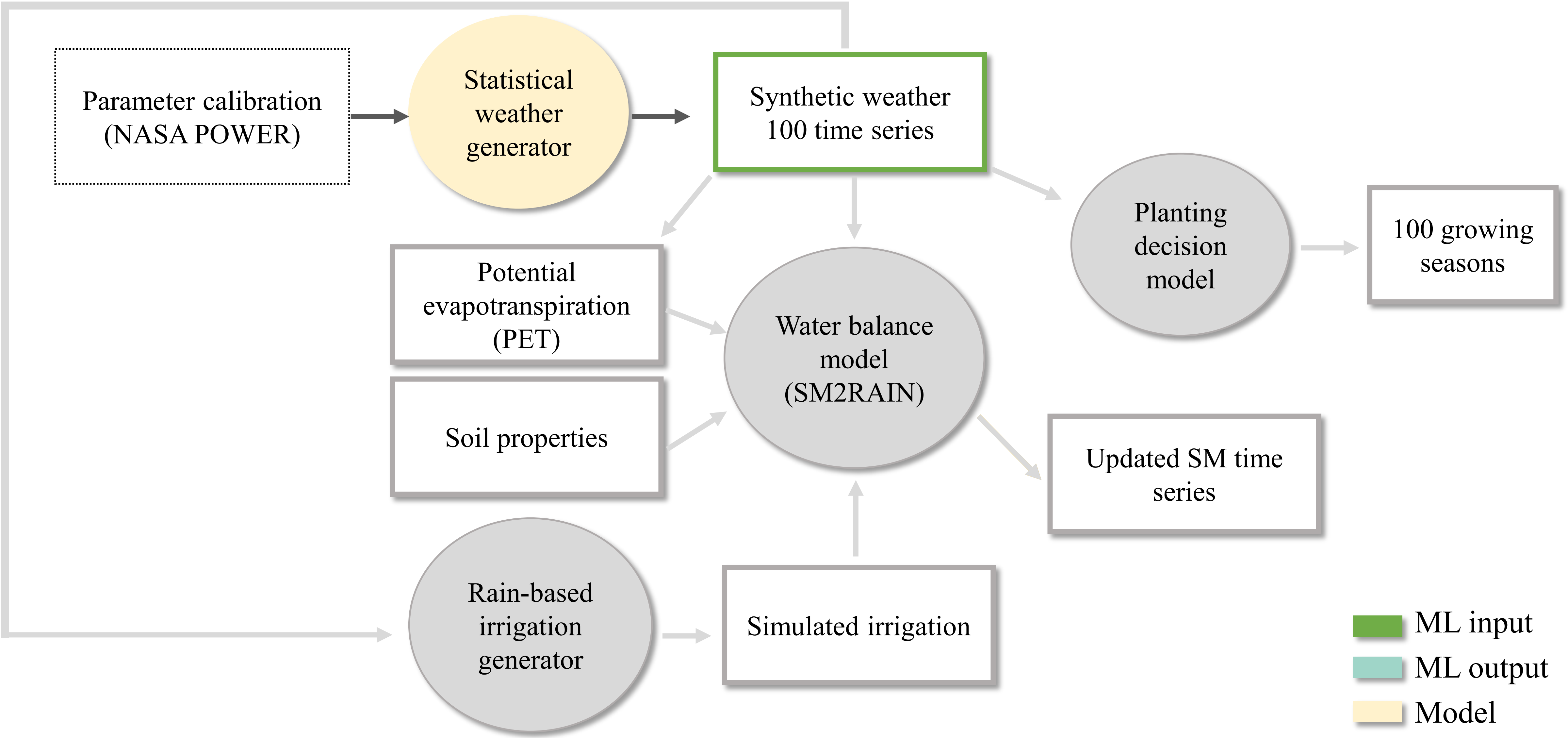
SYNTHETIC DATA GENERATION



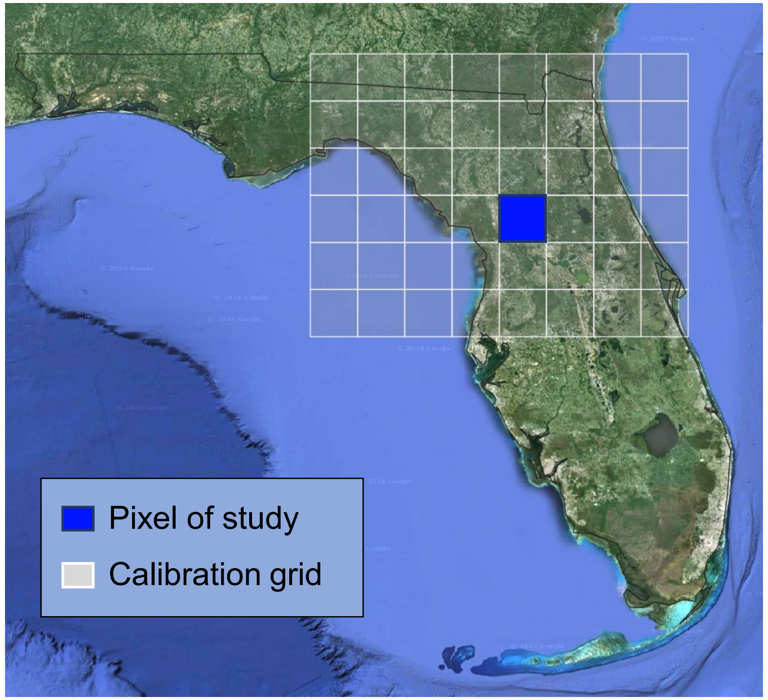
SYNTHETIC DATA GENERATION

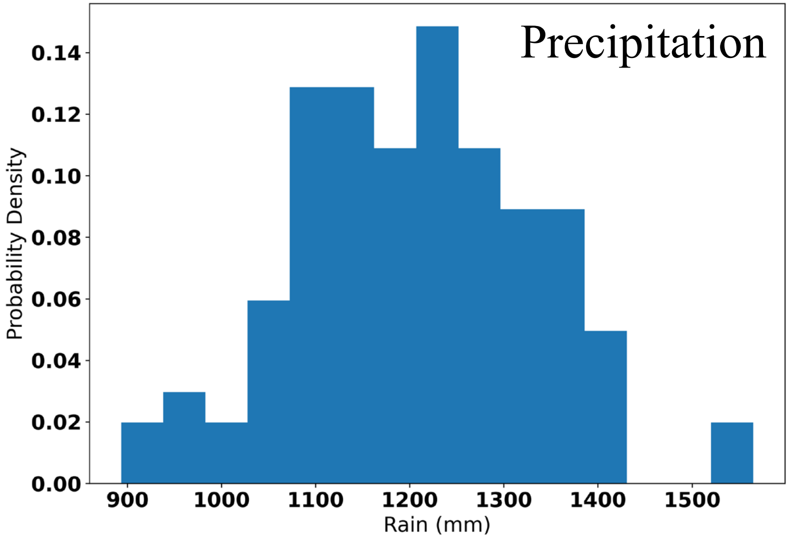


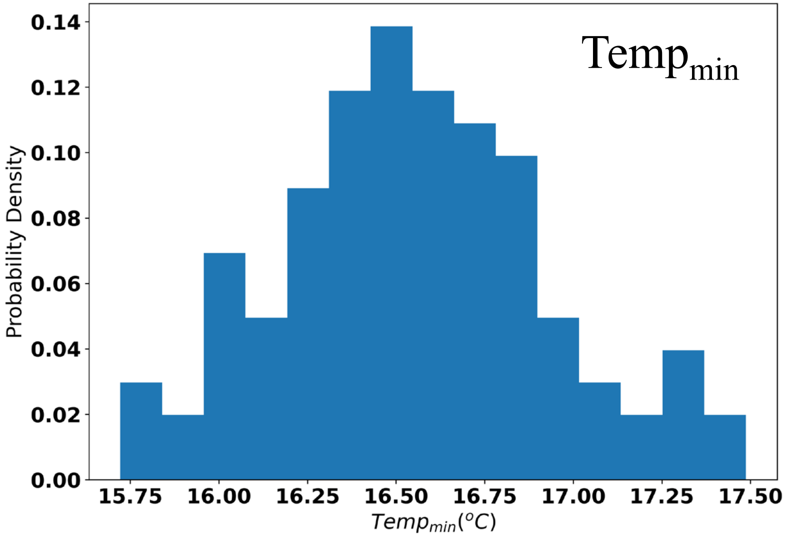
SYNTHETIC DATA GENERATION

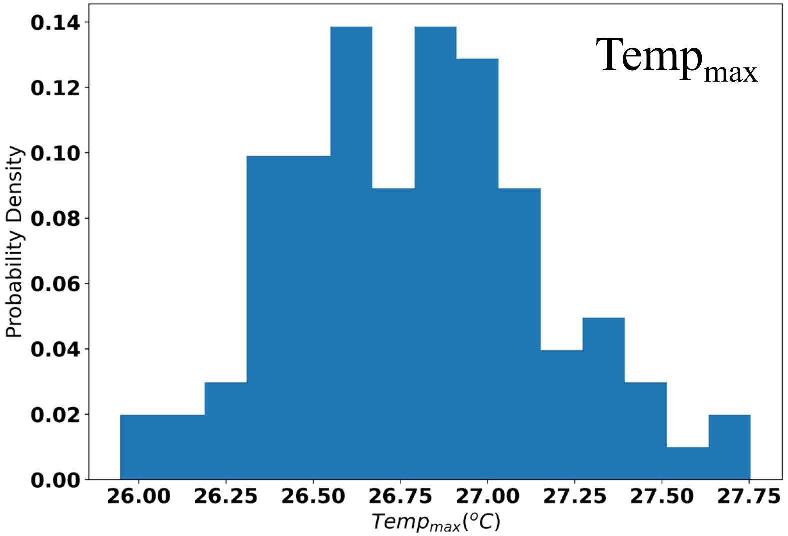


SYNTHETIC WEATHER GENERATOR

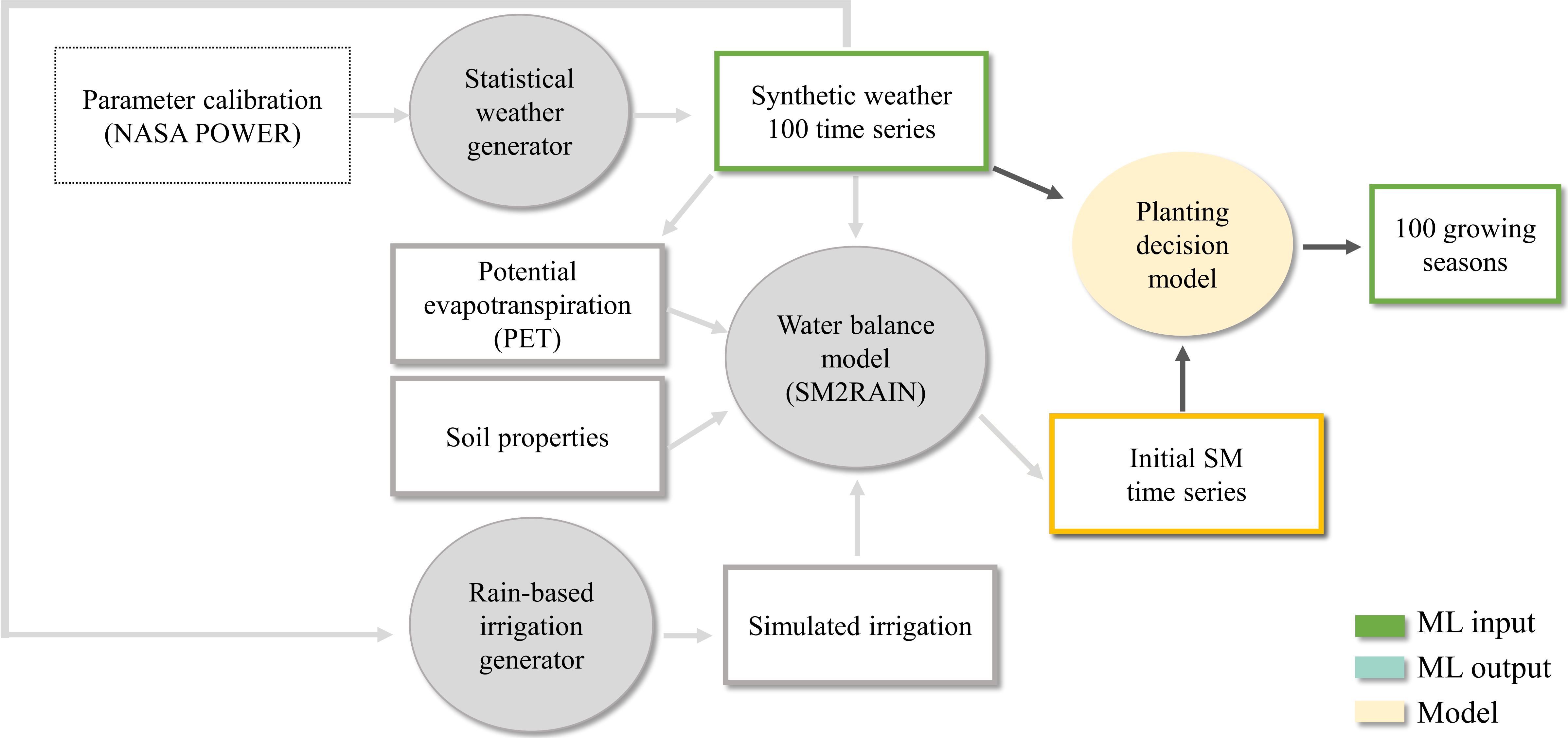








SYNTHETIC DATA GENERATION



PLANTING DECISION MODEL

- Crop: sweet corn → planting period:

Early season or Late season

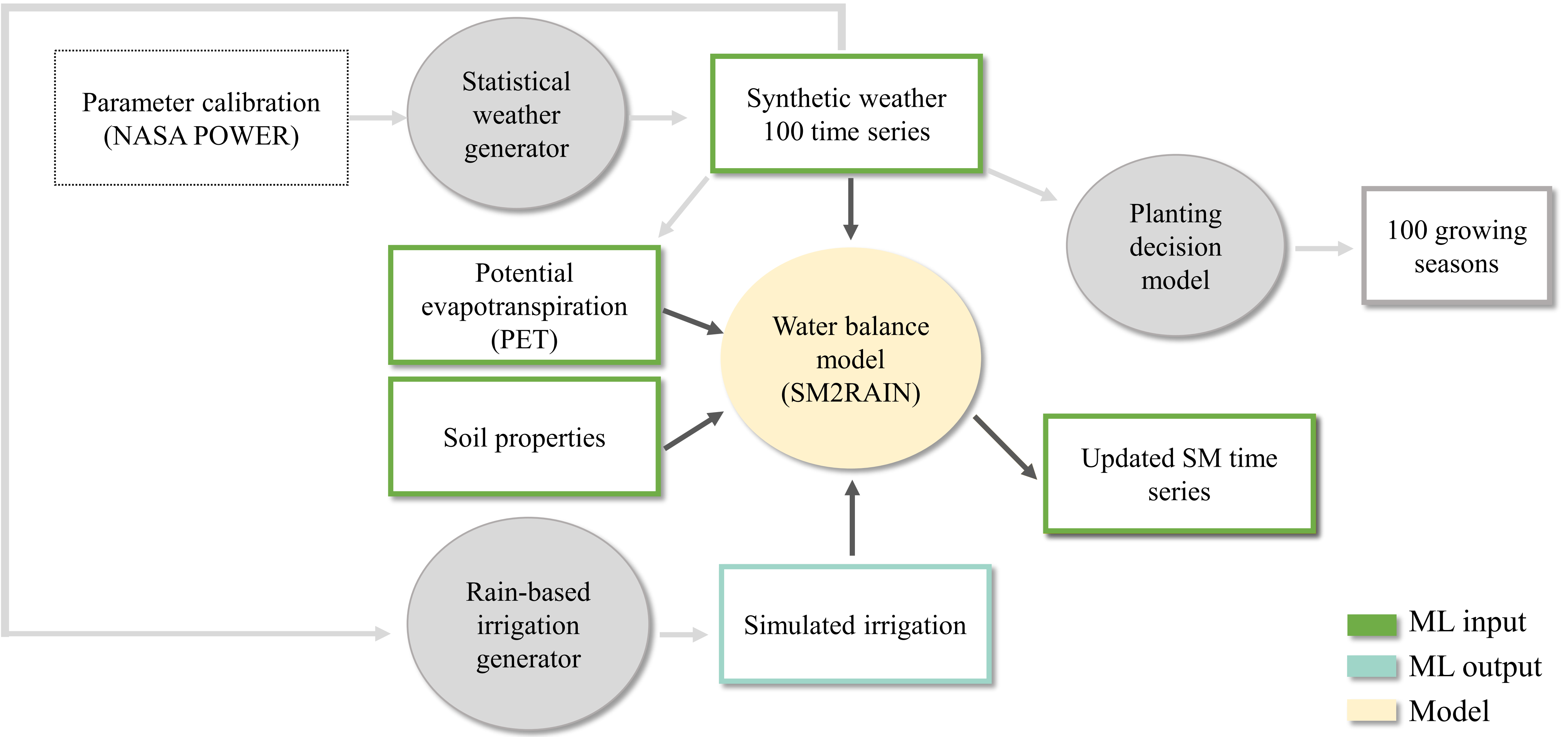
JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC

- Suitable planting days within the planting period:
  - Without precipitation
  - Relative SM range 0.4 - 0.9
  - Minimum temperatures > 7 °C

Day of planting randomly selected

- Length of the season randomly selected from 74 ±4 days

SYNTHETIC DATA GENERATION



WATER BALANCE MODEL (SM2RAIN)

WATER BALANCE

$$nZ\frac{dS(t)}{dt} = r(t) + i(t) - g(t) - e(t)$$

$$\downarrow$$

$$\boxed{nZ(S_t - S_{t-1}) = WI_t - S_t PET_t - K_s {S_t}^{3+\frac{2}{\lambda}}}$$

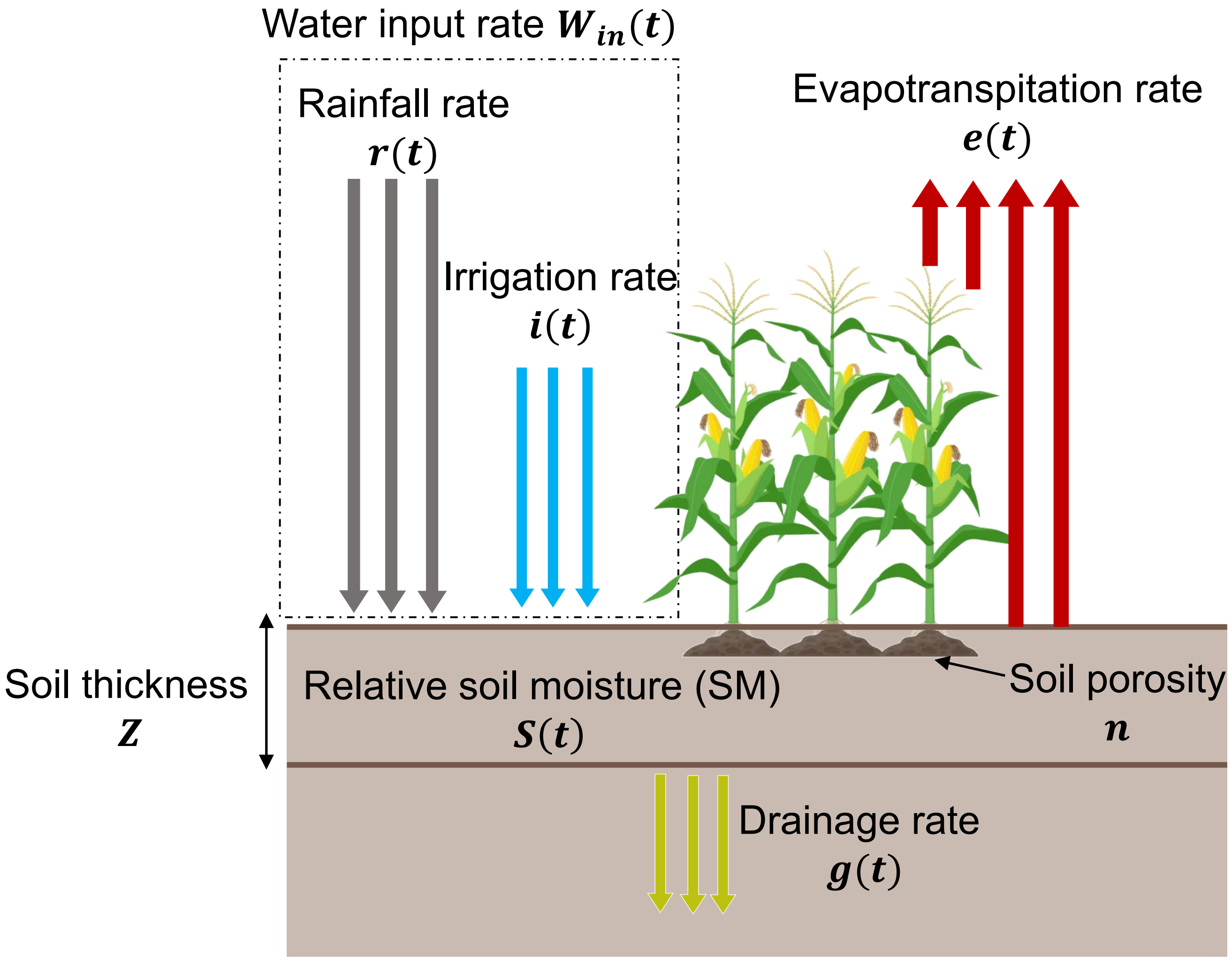
$$S(t) = \left[\frac{\theta(t) - \theta_r}{\theta_s - \theta_r}\right]$$

$\theta(t)$ Volumetric SM
θ_r Residual SM
θ_s Saturation SM

$$e_t = S_t PET_t$$

$$g_t = K_s {S_t}^{3+\frac{2}{\lambda}}$$

K_s Saturated hydraulic conductivity
λ Pore size distribution index



WATER BALANCE MODEL (SM2RAIN)

WATER BALANCE

$$nZ\frac{dS(t)}{dt} = r(t) + i(t) - g(t) - e(t)$$

$$\downarrow$$

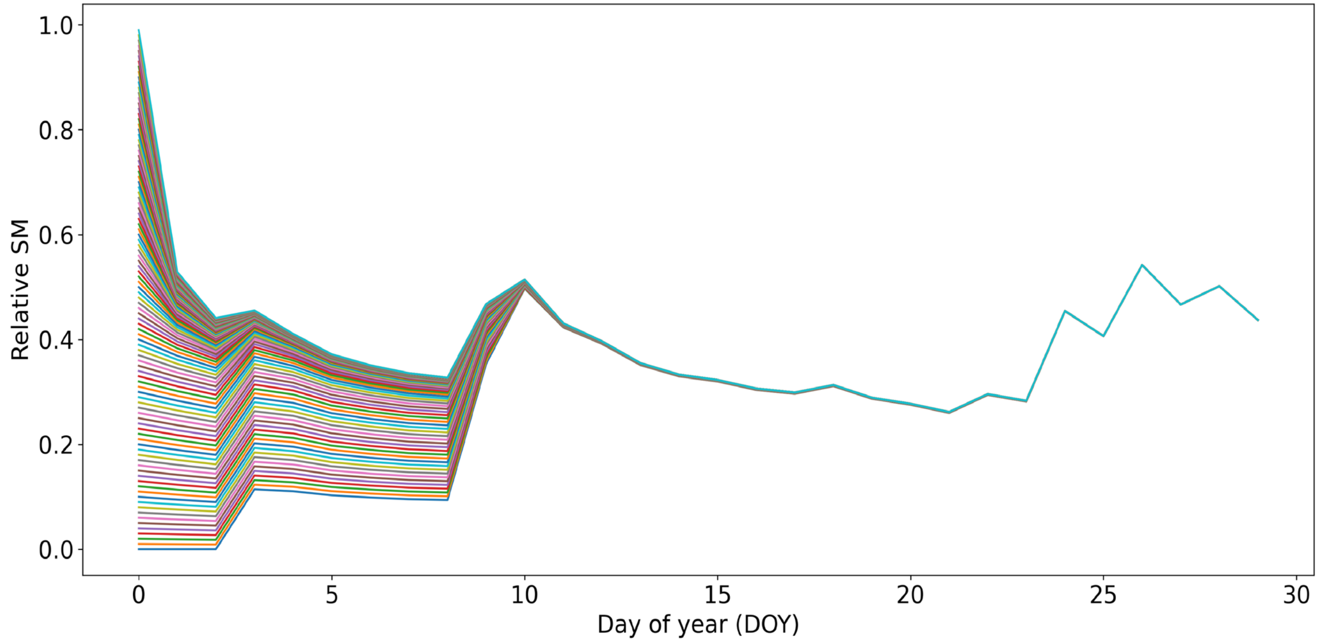
$$\boxed{nZ(S_t - S_{t-1}) = WI_t - S_t PET_t - K_s {S_t}^{3+\frac{2}{\lambda}}}$$

$$K_s {\mathbf{S_t}}^{3+\frac{2}{\lambda}} + (nZ + PET_t)\mathbf{S_t} - (WI_t + nZS_{t-1}) = 0$$

- Equation type to solve knowing S_{t-1}

$$A{\mathbf{S_t}}^{\alpha} + B\mathbf{S_t} - C = 0$$

- Initial value of relative SM ($S_{t=0}$) for Z = 16 cm :



Synthetic time series starts at 12th day.

MACHINE LEARNING FRAMEWORK

- **ML framework A:** Support vector regression (SVR)

$$X = [x_1 \;\; x_2 \;\; \dots \;\; x_n] \qquad \qquad Y$$

Weather, SM, PET → **SVR** → **Irrigation occurrence & amount**

- **ML framework B:** Support vector classification (SVC) + SVR

$$X_1 = [x_1 \;\; x_2 \;\; \dots \;\; x_n] \qquad Y_1 \qquad X_2 = [x_1 \;\; x_2 \;\; \dots \;\; x_{n+1}] \qquad Y_2$$

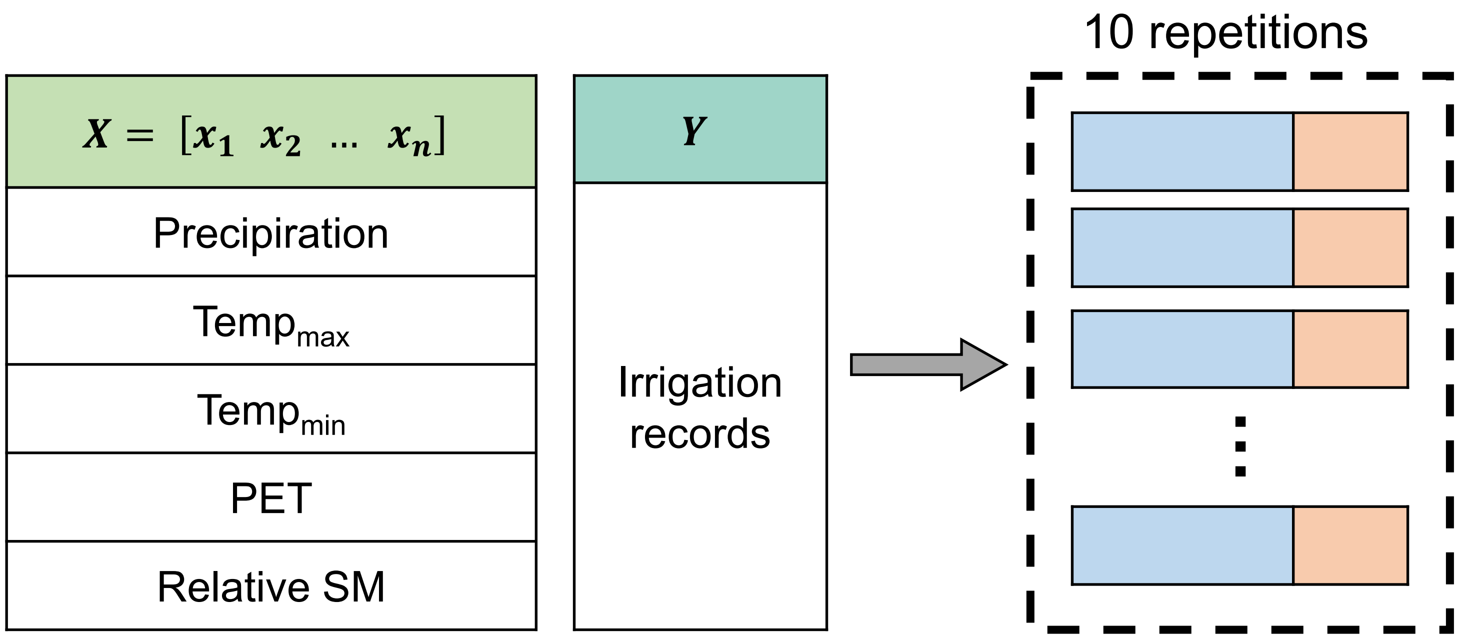
Weather, SM, PET → **SVC** → **Irrigation occurrence** → Weather, SM, PET, Irrigation occurrence → **SVR** → **Irrigation amount**

VALIDATION OF THE ML FRAMEWORK

- **Holdout cross Cross validation:**
  - 100 growing seasons of synthetic data →

set training (70 years)

set testing (30 years)



RESULTS CROSS VALIDATION

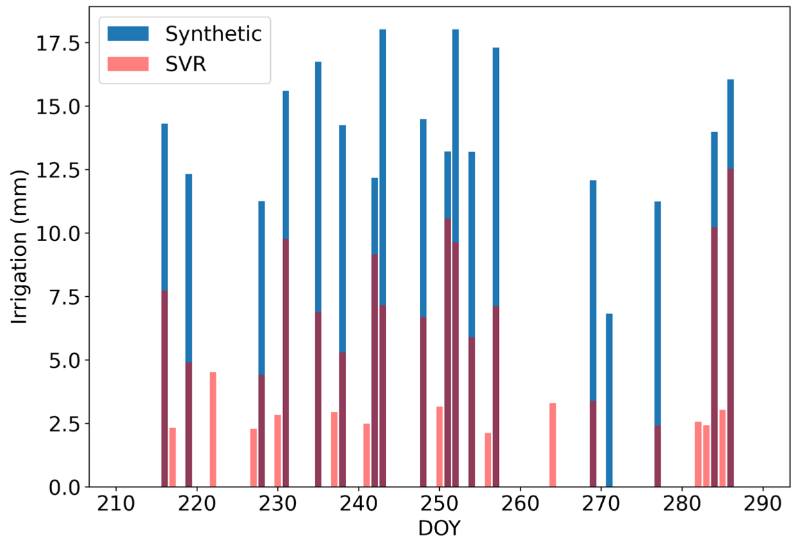
- ML framework A (SVR):

Occurrence

%		Estimated	
		No irr.	Irr.
True	No Irr.	64.57	20.72
	Irr.	0.25	14.46

Amount

RMSE (mm)	6.29
% days underestimating	99.27
% days overestimating	0.73



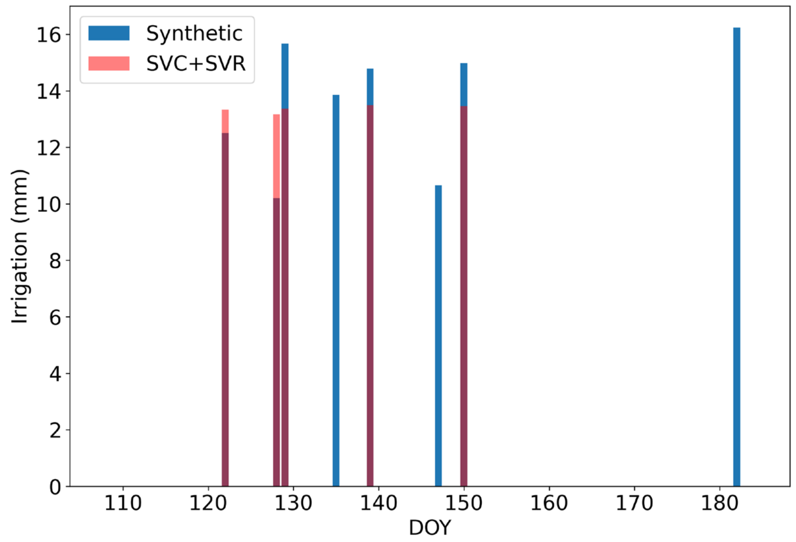
- ML framework B (SVC + SVR):

Occurrence

%		Estimated	
		No irr.	Irr.
True	No Irr.	85.51	0.15
	Irr.	7.36	6.97

Amount

RMSE (mm)	2.01
% days underestimating	60.90
% days overestimating	39.10



SUMMARY AND FUTURE WORK

Summary:

- Applying SVR leads to underestimation of irrigation.
- SVR tends to detect more false events than SVC, while SVC misses more events.
- The amount of irrigation can be better estimated with information about irrigation occurrence.

Future work:

- Use neural networks (NN) that account for more complex relationships.
- Evaluate the framework with *in situ* data.

THANK YOU

FUNDINGS SUPPORT FROM USDA-CIG AWARD #P0206601

WATER BALANCE (SM2RAIN)

$$nZ\frac{dS(t)}{dt} = r(t) + i(t) - e(t) - g(t) - \cancelto{0}{sr(t)}$$

$$\boxed{e_t = S_t PET_t} \quad \boxed{g_t = K_s {S_t}^{3+\frac{2}{\lambda}}} \quad \boxed{WI_t = r_t + i_t}$$

$$nZ(S_t - S_{t-1}) = WI_t - S_t PET_t - K_s {S_t}^{3+\frac{2}{\lambda}}$$

$$K_s {S_t}^{3+\frac{2}{\lambda}} + (nZ + PET_t)S_t - (WI_t + nZS_{t-1}) = 0$$

$$\boxed{A \equiv K_s} \quad \boxed{B \equiv nZ + PET_t S_t} \quad \boxed{C \equiv WI_t + nZS_{t-1}} \quad \boxed{\alpha \equiv 3 + \frac{2}{\lambda}}$$

$$A{S_t}^{\alpha} + BS_t - C = 0$$

Newton Raphson Method

$$f(S_t) = A{S_t}^{\alpha} + BS_t - C$$

$$f'(S_t) = \alpha A{S_t}^{\alpha-1} + B$$

$$S_t^{n+1} = S_t^n - \frac{f(S_t^n)}{f'(S_t^n)}$$

Start $S_t^0 = S_{t-1}$

End $S_t = S_t^{n+1}$ when $|S_t^n - S_t^{n+1}| < 0.001$