Merging the Gap between Data Driven Approach and its Field Application: a Transmission Line Fault Location Study

Yu Liu, Associate Professor
ShanghaiTech University

Email: liuyu.shanghaitech@gmail.com
liuyu@shanghaitech.edu.cn

July 20, 2023

Power System Protection and Automation Laboratory
电力系统保护与自动化实验室

Outline

- ➢ Challenges of Data Driven Approaches for Fault Location
- ➢ Merging the Gap: Physics-Informed Data Driven Method
- ➢ Field Data Results
- ➢ Discussion
- ➢ Conclusion & Future Work

What is Fault Location?

- **Fault occurs in transmission line**
- **Fault is isolated by the circuit breakers controlled by protective relays**
- **Estimate the location of the fault within the line of interest using the measurements at line terminals before and during the fault**

Existing methods:

- **Model based method (using phasors)**

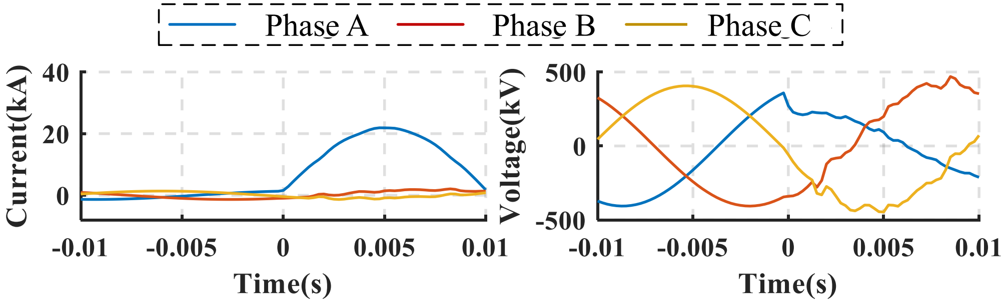
– Key Issue: Require long time window during faults (typically more than 1 cycle)

- **Traveling wave based method**

– Key Issue: Extremely high sampling rate (typically in the order of MHz)

- **Data-driven based method**

– Key Issue: Availability of High Quality Data



Fault Location

Challenges of Data Driven Approaches for Fault Related Applications

Data-driven based method

– Key Issue: Availability of High Quality Data

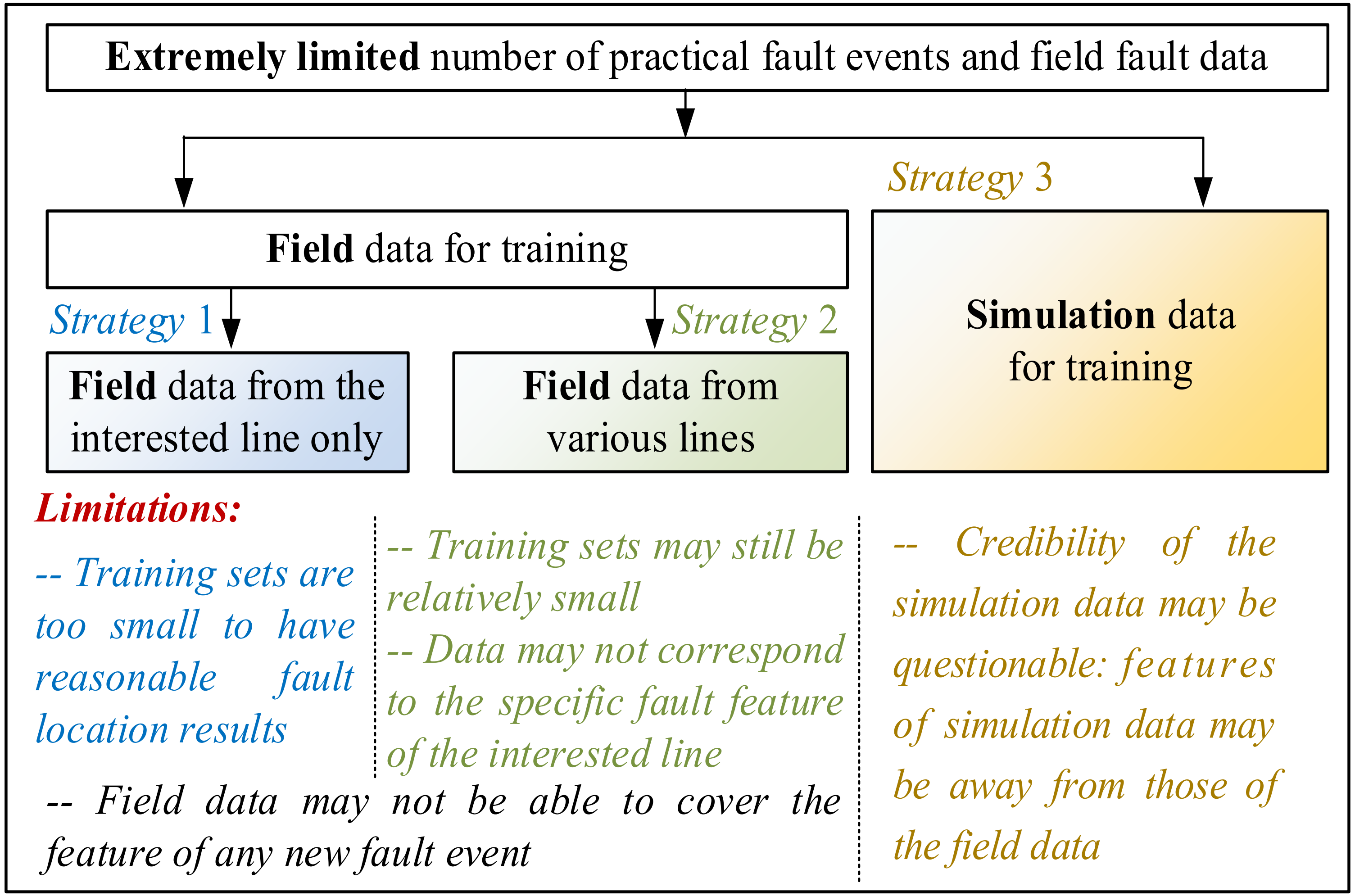
- **Field Data during normal operation:** Extensive
- **Field Data during faults: Extremely Limited**

Facts:

Transmission lines (≥ 220kV) in State Grid Corporation of China:

- **Overall line length 6.2×10^5 km**
- **2000 faults in year 2020 [1].**

- **On average: for a 310 km line, only 1 fault/year**



[1] Annual National Reliability Report of China, National Energy Administration, 2020.

Challenges of Data Driven Approaches for Fault Related Applications

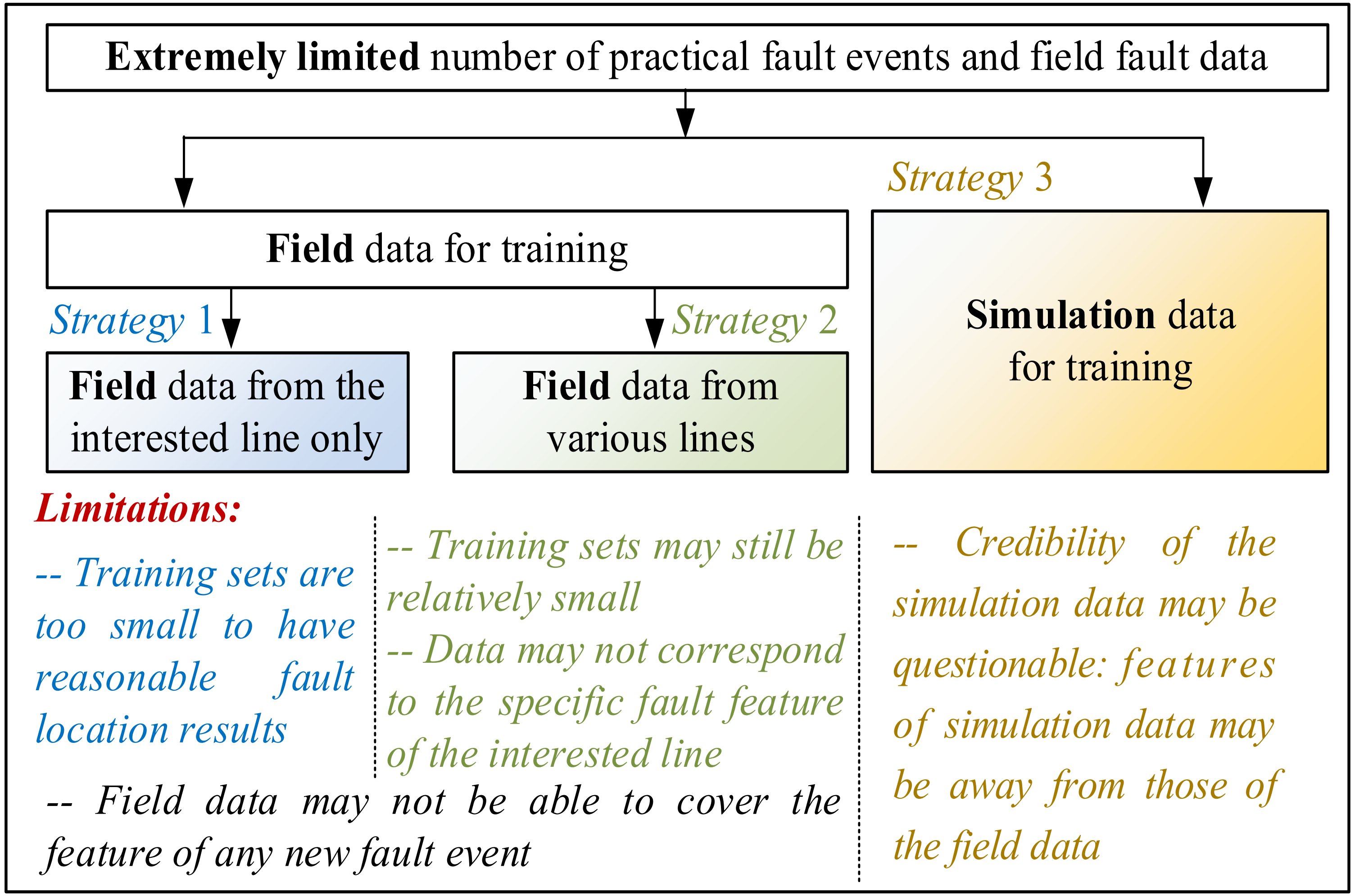
- **Key Features to solve the above challenges?**

- **Key Idea:**

1. Only field data for training is not enough!! One must use simulation data for training.

2. To ensure practicability, one must use field data for testing!

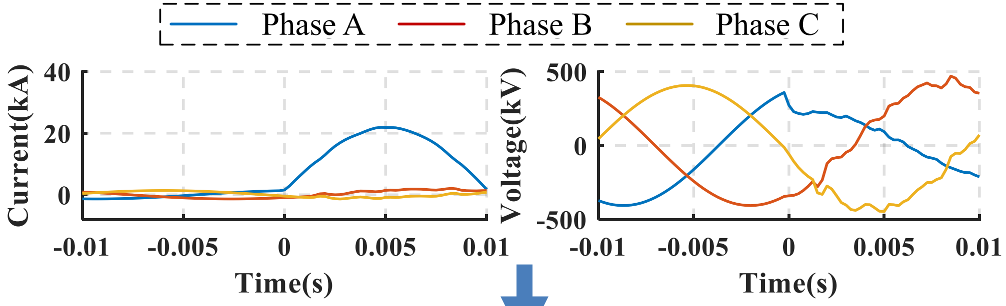
3. We have to Merge the GAP between Simulation and Field Data!



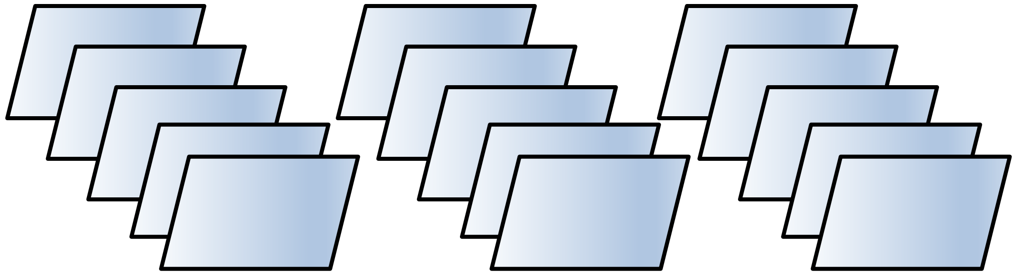
Challenges of Data Driven Approaches for Fault Related Applications

We have to Merge the GAP between Simulation and Field Data!

Field Data of 1 Fault Event



Generate the "target dataset" via Simulation including fault events



Idea of the Proposed Data Driven Fault Location

Start

Step 1

Estimate the ranges of system parameters
(source information, loading condition, FIAs and fault resistances)
using single-ended measured field data and physical laws

Step 2

Generate the "target dataset" using EMT simulation tools
(consistent with the fault features and format of the field data)
to greatly augment the dataset for training

Step 3

Train the neural network using the "target dataset"
to achieve fault location compatible with field data

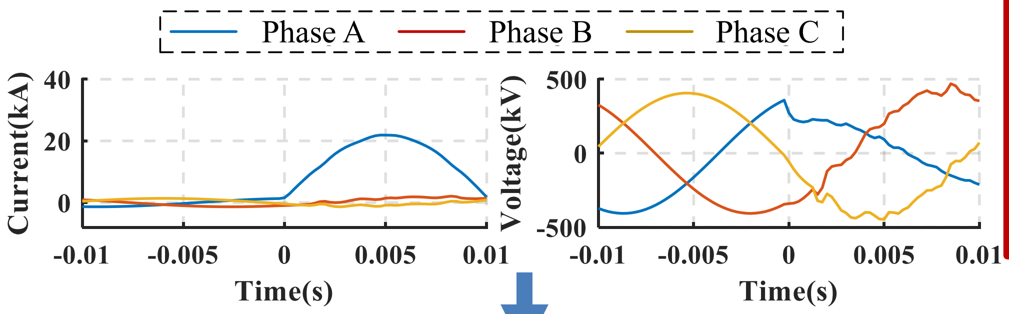
Outline

- Challenges of Data Driven Approaches for Fault Location
- Merging the Gap: Physics-Informed Data Driven Method
- Field Data Results
- Discussion
- Conclusion & Future Work

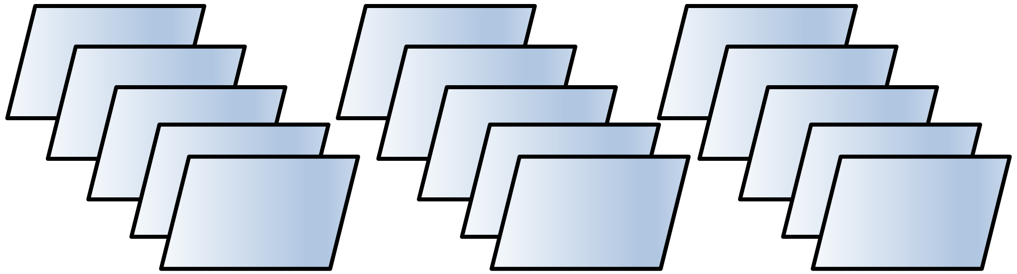
Physics-Informed Data Driven Method

We have to Merge the GAP between Simulation and Field Data!

Field Data of 1 Fault Event



Generate the "target dataset" via Simulation including fault events



Idea of the Proposed Data Driven Fault Location

Start

Step 1

Estimate the ranges of system parameters
(source information, loading condition, FIAs and fault resistances)
using single-ended measured field data and physical laws

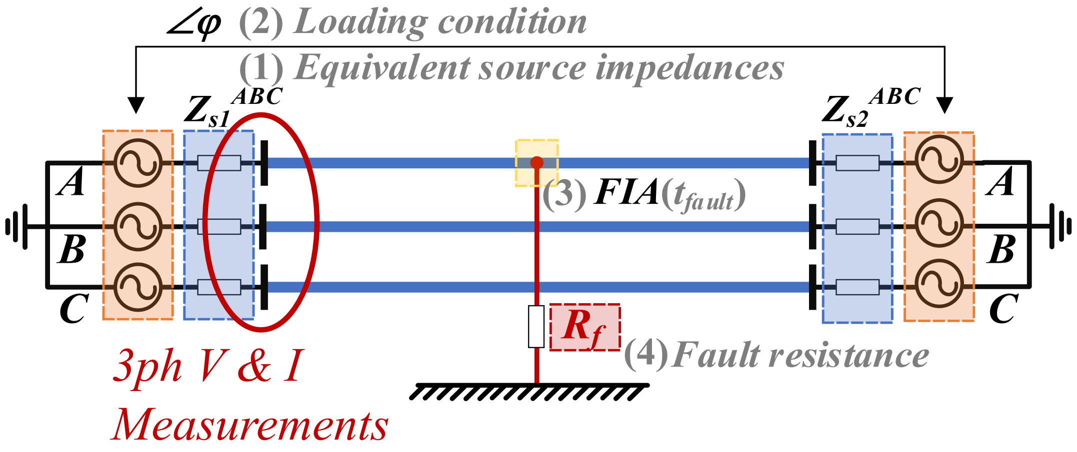
Step 2

Generate the "target dataset" using EMT simulation tools
(consistent with the fault features and format of the field data)
to greatly augment the dataset for training

Step 3

Train the neural network using the "target dataset"
to achieve fault location compatible with field data

Step 1: Estimate system parameters (Physics)



(0) Preparations

Carried out in the $\alpha\beta0$ mode of the system

$$\begin{cases} \boldsymbol{u}^{\alpha\beta0} = \boldsymbol{T}_{Clarke}^{-1}\boldsymbol{u}^{ABC} \\ \boldsymbol{i}^{\alpha\beta0} = \boldsymbol{T}_{Clarke}^{-1}\boldsymbol{i}^{ABC} \end{cases}, \begin{cases} \boldsymbol{R}^{\alpha\beta0} = \boldsymbol{T}_{Clarke}^{-1}\boldsymbol{R}^{ABC}\boldsymbol{T}_{Clarke} \\ \boldsymbol{L}^{\alpha\beta0} = \boldsymbol{T}_{Clarke}^{-1}\boldsymbol{L}^{ABC}\boldsymbol{T}_{Clarke} \end{cases}$$

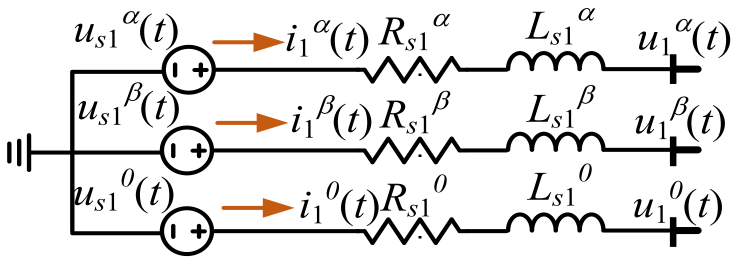
$$\boldsymbol{T}_{Clarke}^{-1} = [2/3,-1/3,-1/3;0,1/\sqrt{3},-1/\sqrt{3};1/3,1/3,1/3]$$

Calculate initiation time t_f

$$t_f = t\,|_{i_1^{j(1)} > k_{tf} I_1^{j(1)}{}_{\max}}$$

(1) Equivalent source impedance

Equivalent source model for three phase transmission lines



$$u_{s1}^{j}(t) - u_1^{j}(t) = R_{s1}^{j} i_1^{j}(t) + L_{s1}^{j} di_1^{j}(t)/dt \quad \text{During the fault} \quad (1)$$

$$u_{s1}^{j}(t-\Delta T) - u_1^{j}(t-\Delta T) = R_{s1}^{j} i_1^{j}(t-\Delta T) + L_{s1}^{j} di_1^{j}(t-\Delta T)/dt \quad \text{Before the fault} \quad (2)$$

(1) - (2): $u_1^{j}(t) - u_1^{j}(t-\Delta T) = R_{s1}^{j}\left[i_1^{j}(t) - i_1^{j}(t-\Delta T)\right] + L_{s1}^{j} d\left[i_1^{j}(t) - i_1^{j}(t-\Delta T)\right]/dt$

Solved with the least square scheme

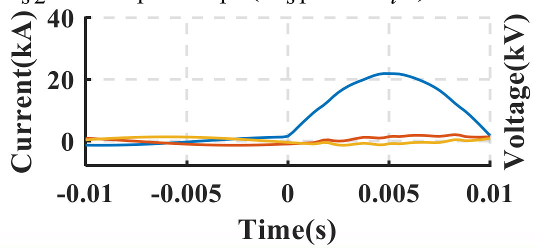
$$\left[\Delta \boldsymbol{i}_1^{j} \quad \Delta \boldsymbol{i}_1^{j(1)}\right]\boldsymbol{x} = \left[\Delta \boldsymbol{u}_1^{j}\right] \longrightarrow \boldsymbol{x} = \left(\left[\Delta \boldsymbol{i}_1^{j} \quad \Delta \boldsymbol{i}_1^{j(1)}\right]^T\left[\Delta \boldsymbol{i}_1^{j} \quad \Delta \boldsymbol{i}_1^{j(1)}\right]\right)^{-1}\left[\Delta \boldsymbol{i}_1^{j} \quad \Delta \boldsymbol{i}_1^{j(1)}\right]^T\left[\Delta \boldsymbol{u}_1^{j}\right]$$

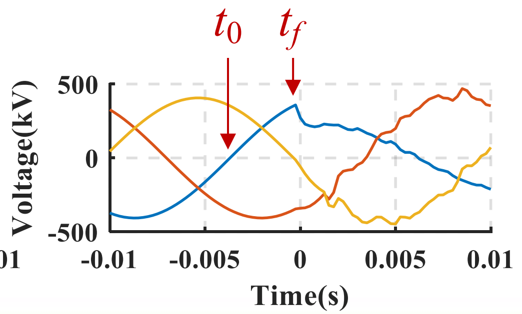
(2) Loading condition (φ_2-φ_1)

$$\begin{cases} \tilde{U}_{s1}^{j} = \tilde{U}_1^{j} + \tilde{I}_1^{j} Z_{s1}^{j} \\ \tilde{U}_{s2}^{j} = \tilde{U}_1^{j} - \tilde{I}_1^{j}(Z_{s1}^{j} + Z_l^{j}) \end{cases}$$

(3) Fault inception angle

$$FIA = 2\pi(t_f - t_0)/\Delta T$$





Step 1: Estimate system parameters (Physics)

(4) Fault resistance R_f (admittance Y_f)

Key Idea: With guessed R_f and l_f, one can calculate $i_1^j(t)$; compare it with the actual current measurement to get range of R_f.

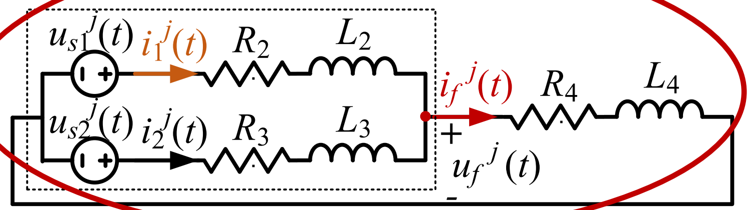
	SLG (AG)	LL (BC)	LLG (BCG)	3PH (ABC)
The fault matrix $\boldsymbol{Y}_f^{\alpha\beta 0}$	$\begin{bmatrix} \frac{2}{3}Y_f & 0 & \frac{2}{3}Y_f \\ 0 & 0 & 0 \\ \frac{1}{3}Y_f & 0 & \frac{1}{3}Y_f \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2Y_f & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{3}Y_f & 0 & -\frac{2}{3}Y_f \\ 0 & Y_f & 0 \\ -\frac{1}{3}Y_f & 0 & \frac{2}{3}Y_f \end{bmatrix}$	$\begin{bmatrix} 3Y_f & 0 & 0 \\ 0 & 3Y_f & 0 \\ 0 & 0 & 0 \end{bmatrix}$

During LL, LLG, 3PH faults:

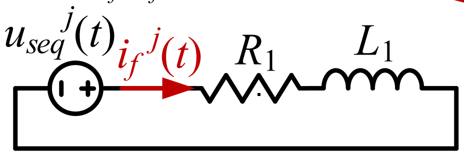
$$u_{s1}^{\beta}(t) - u_f^{\beta}(t) = R_{eq1}^{\beta} i_1^{\beta}(t) + L_{eq1}^{\beta}\, di_1^{\beta}(t)/dt$$

$$u_{s2}^{\beta}(t) - u_f^{\beta}(t) = R_{eq2}^{\beta} i_2^{\beta}(t) + L_{eq2}^{\beta}\, di_2^{\beta}(t)/dt$$

$$i_1^{\beta}(t) + i_2^{\beta}(t) = i_f^{\beta}(t) \qquad i_f^{\beta}(t) = 2Y_f u_f^{\beta}(t)$$



$$di_f^{\,j}(t)/dt + B_1 i_f^{\,j}(t) = B_2(t)$$

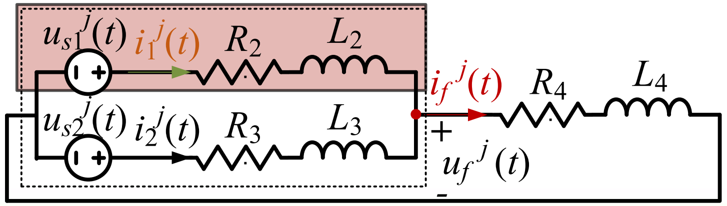


Simplify to

Recalculate $i_a^j(t)$

$$di_1^{\,j}(t)/dt + B_3 i_1^{\,j}(t) = B_4(t)$$

$$i_1^{\,j}(t)$$



During SLG faults:

Both α mode and 0 mode

$$u_{s1}^{\alpha}(t) - u_f^{\alpha}(t) = R_{eq1}^{\alpha} i_1^{\alpha}(t) + L_{eq1}^{\alpha}\, di_1^{\alpha}(t)/dt$$

$$u_{s2}^{\alpha}(t) - u_f^{\alpha}(t) = R_{eq2}^{\alpha} i_2^{\alpha}(t) + L_{eq2}^{\alpha}\, di_2^{\alpha}(t)/dt$$

$$-u_f^{0}(t) = R_{eq1}^{0} i_1^{0}(t) + L_{eq1}^{0}\, di_1^{0}(t)/dt$$

$$-u_f^{0}(t) = R_{eq2}^{0} i_2^{0}(t) + L_{eq2}^{0}\, di_2^{0}(t)/dt$$

$$i_f^{\alpha}(t) = 2Y_f/3 \cdot u_f^{\alpha}(t) + 2Y_f/3 \cdot u_f^{0}(t)$$

$$i_f^{0}(t) = Y_f/3 \cdot u_f^{\alpha}(t) + Y_f/3 \cdot u_f^{0}(t)$$

$$i_f^{\alpha}(t) = i_1^{\alpha}(t) + i_2^{\alpha}(t) \qquad i_f^{0}(t) = i_1^{0}(t) + i_2^{0}(t)$$

Simplify to α mode

$$u_{s1}^{\alpha}(t) - u_f^{\alpha}(t) = R_{eq1}^{\alpha} i_1^{\alpha}(t) + L_{eq1}^{\alpha}\, di_1^{\alpha}(t)/dt$$

$$u_{s2}^{\alpha}(t) - u_f^{\alpha}(t) = R_{eq2}^{\alpha} i_2^{\alpha}(t) + L_{eq2}^{\alpha}\, di_2^{\alpha}(t)/dt$$

$$u_f^{\alpha}(t) = 1/2 \cdot R_{eq}^{0} i_f^{\alpha}(t) + 1/2 \cdot L_{eq}^{0}\, di_{f\alpha}(t)/dt + 3/2Y_f \cdot i_f^{\alpha}(t)$$

$$i_f^{\alpha}(t) = i_1^{\alpha}(t) + i_2^{\alpha}(t)$$

Step 1: Estimate system parameters (Physics)

(4) Fault resistance R_f (admittance Y_f)

$$i_f^{\,j}(t) = -M_{sol1}\cos(\varphi_{sol1})e^{-B_1 t} + M_{sol1}\cos(\omega t + \varphi_{sol1})$$

$$i_1^{\,j}(t) = [i_1^{\,j}(0_+) - M_{sol2}\cos(\varphi_{sol2}) - A_{step2}/(B_3 - B_1)]e^{-B_3 t}$$
$$+M_{sol2}\cos(\omega t + \varphi_{sol2}) + A_{step2}/(B_3 - B_1)e^{-B_1 t}$$

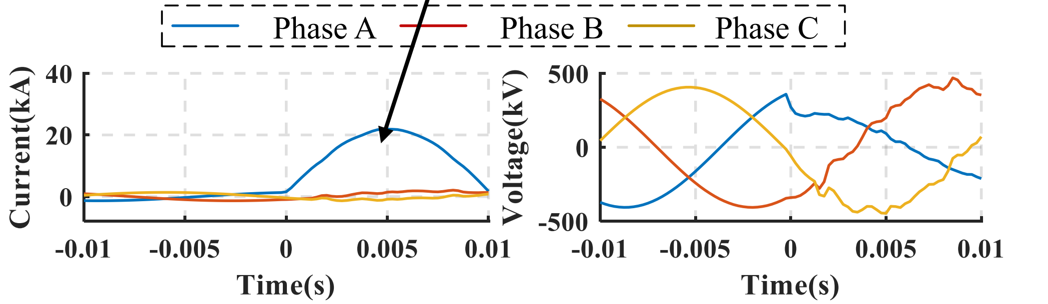
*$i_f^{\,j}(t)$ is a function of **fault resistance R_f** and **fault location l_f***

Type	SLG	LL	LLG	3PH
$u_{seq}^{\,j}$	$u_{seq}^{\,\alpha}$	$u_{seq}^{\,\beta}$	$u_{seq}^{\,\beta}$	$u_{seq}^{\,\beta}$
$u_{s1}^{\,j}$	$u_{s1}^{\,\alpha}$	$u_{s1}^{\,\beta}$	$u_{s1}^{\,\beta}$	$u_{s1}^{\,\beta}$
R_1	$R_{eq}^{\,\alpha}+1/2R_{eq}^{\,0}+3/2Y_f$	$R_{eq}^{\,\beta}+1/2Y_f$	$R_{eq}^{\,\beta}+1/Y_f$	$R_{eq}^{\,\beta}+1/3Y_f$
L_1	$L_{eq}^{\,\alpha}+1/2L_{eq}^{\,0}$	$L_{eq}^{\,\beta}$	$L_{eq}^{\,\beta}$	$L_{eq}^{\,\beta}$
R_2	$R_{eq1}^{\,\alpha}$	$R_{eq1}^{\,\beta}$	$R_{eq1}^{\,\beta}$	$R_{eq1}^{\,\beta}$
L_2	$L_{eq1}^{\,\alpha}$	$L_{eq1}^{\,\beta}$	$L_{eq1}^{\,\beta}$	$L_{eq1}^{\,\beta}$
R_4	$1/2R_{eq}^{\,0}+3/2Y_f$	$1/2Y_f$	$1/Y_f$	$1/3Y_f$
L_4	$1/2L_{eq}^{\,0}$	0	0	0

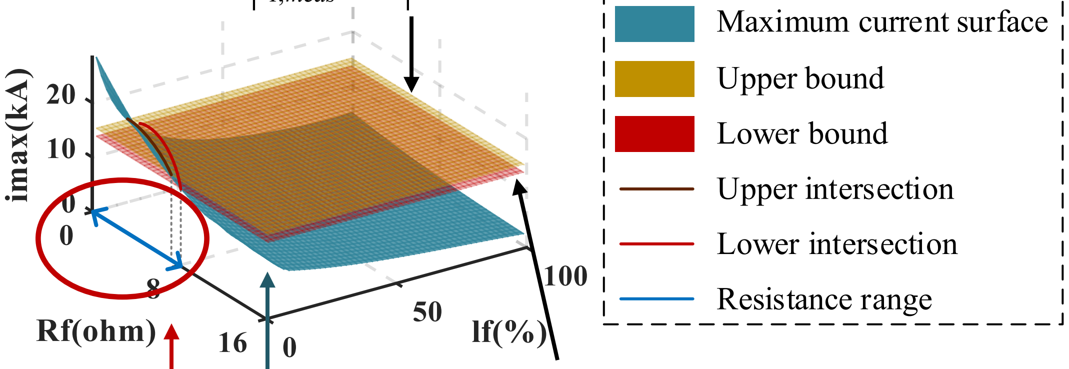
Criteria for determining fault resistance (typical value of c: 5%)

$$(1-c)\cdot\max\left|i_{1,meas}^{\,j}(t)\right| \le \max_{R_f}\left|i_1^{\,j}(t)\right| \le (1+c)\cdot\max\left|i_{1,meas}^{\,j}(t)\right|$$

Eg. *Max current = 21.09 kA,*
Max current of alpha mode $i_{1,meas}^{\,\alpha}$ = 14.63 kA



$$(1+c)\cdot\max\left|i_{1,meas}^{\,\alpha}(t)\right| = 15.36\,kA$$



$$(1-c)\cdot\max\left|i_{1,meas}^{\,\alpha}(t)\right| = 13.90\,kA$$

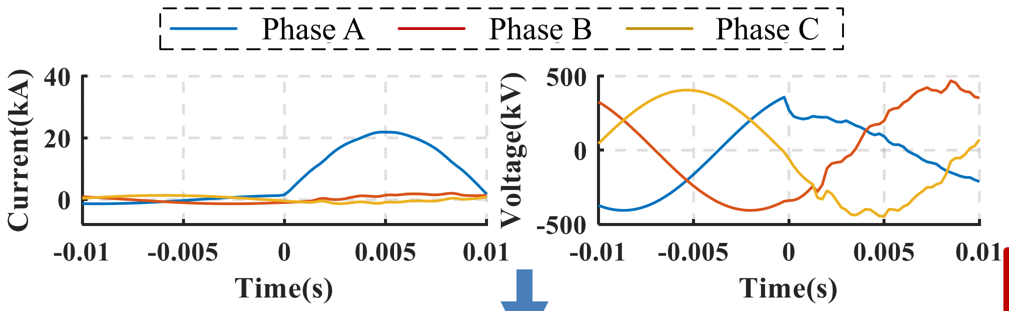
Range of R_f

The maximum current surface obtained through formula derivation: $i_f^{\,j}(t)$ as functions of R_f and l_f

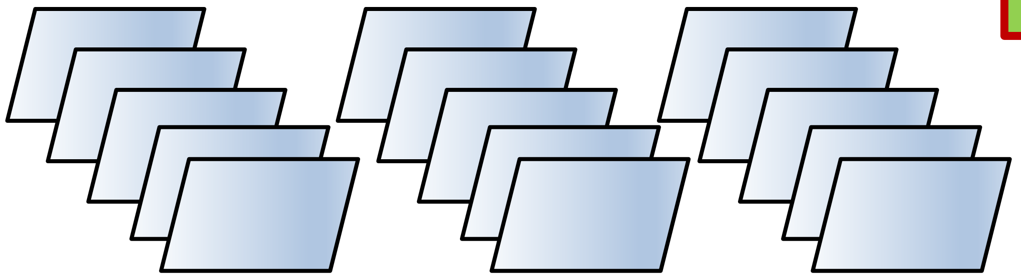
Physics-Informed Data Driven Method

We have to Merge the GAP between Simulation and Field Data!

Field Data of 1 Fault Event



Generate the "target dataset" via Simulation including fault events



Idea of the Proposed Data Driven Fault Location

Start

Step 1

Estimate the ranges of system parameters
(source information, loading condition, FIAs and fault resistances)
using single-ended measured field data and physical laws

Step 2

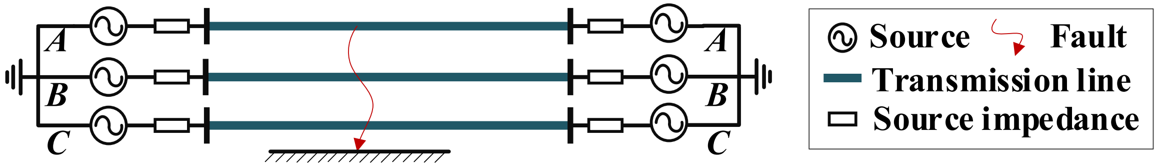
Generate the "target dataset" using EMT simulation tools
(consistent with the fault features and format of the field data)
to greatly augment the dataset for training

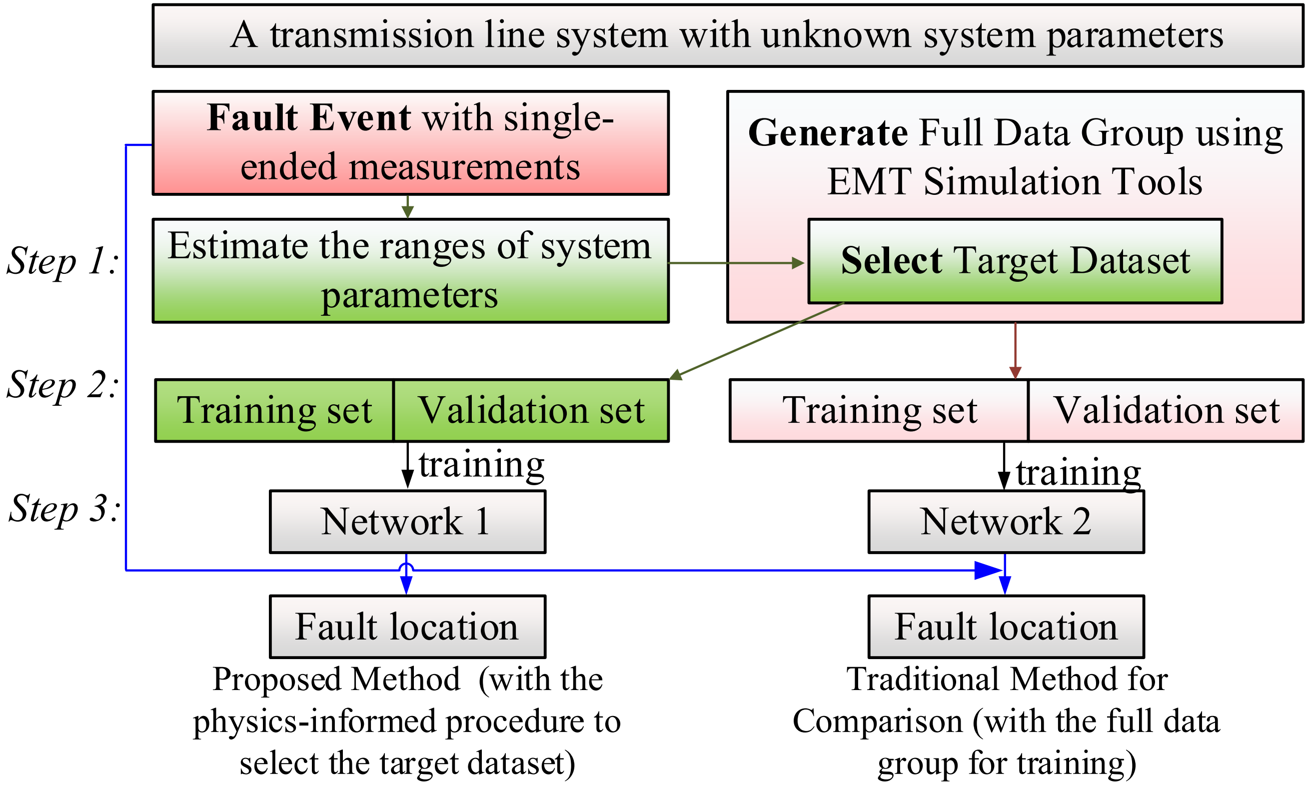
Step 3

Train the neural network using the "target dataset"
to achieve fault location compatible with field data

Step 2: "Target dataset" generation via simulation

Three phase transmission line simulation model





Flow chart of the proposed method

Flow chart of the traditional data-driven method for comparison

Eg. Simulation full dataset and testing dataset

Parameters in the **full data group**, simulation

Parameters	Value
Source impedance/ohm	Z_α=Z_β=0.1+j1, 0.4+j1.5, 1+j5 Z_0=0.2+j1.5, 0.8+j3, 2+j6
Loading condition/deg	±2, ±6, ±10, ±12
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10, 20, 50, 100, 150, 200
Fault location/km	1, 2, 3, …,
FIA/deg	0, 45, 90, …, 270, 315 (step: 45 deg)

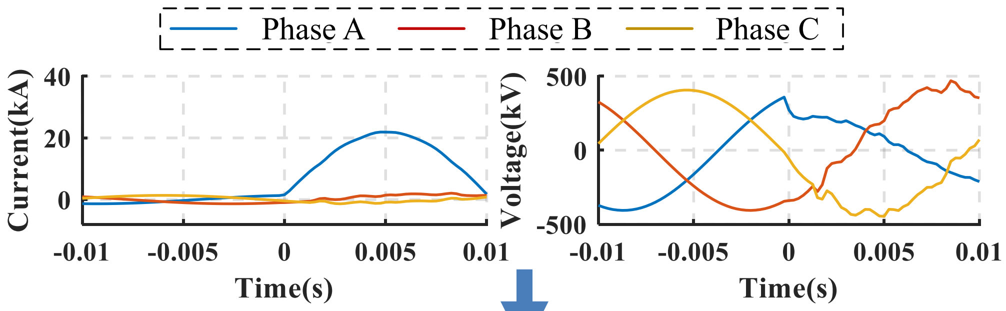
Parameters in the **target dataset**, simulation

Parameters	Value
Source impedance/ohm	Z_α=Z_β= 0.4+j1.5, 1+j5; Z_0=0.8+j3, 2+j6
Loading condition/deg	±2
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10
Fault location/km	1, 2, 3, …, 22
FIA/deg	45, 90

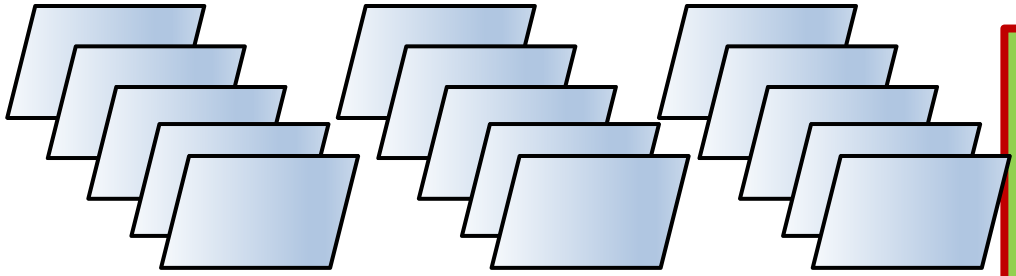
Physics-Informed Data Driven Method

We have to Merge the GAP between Simulation and Field Data!

Field Data of 1 Fault Event



Generate the "target dataset" via Simulation including fault events



Idea of the Proposed Data Driven Fault Location

Start

Step 1

Estimate the ranges of system parameters
(source information, loading condition, FIAs and fault resistances)
using single-ended measured field data and physical laws

Step 2

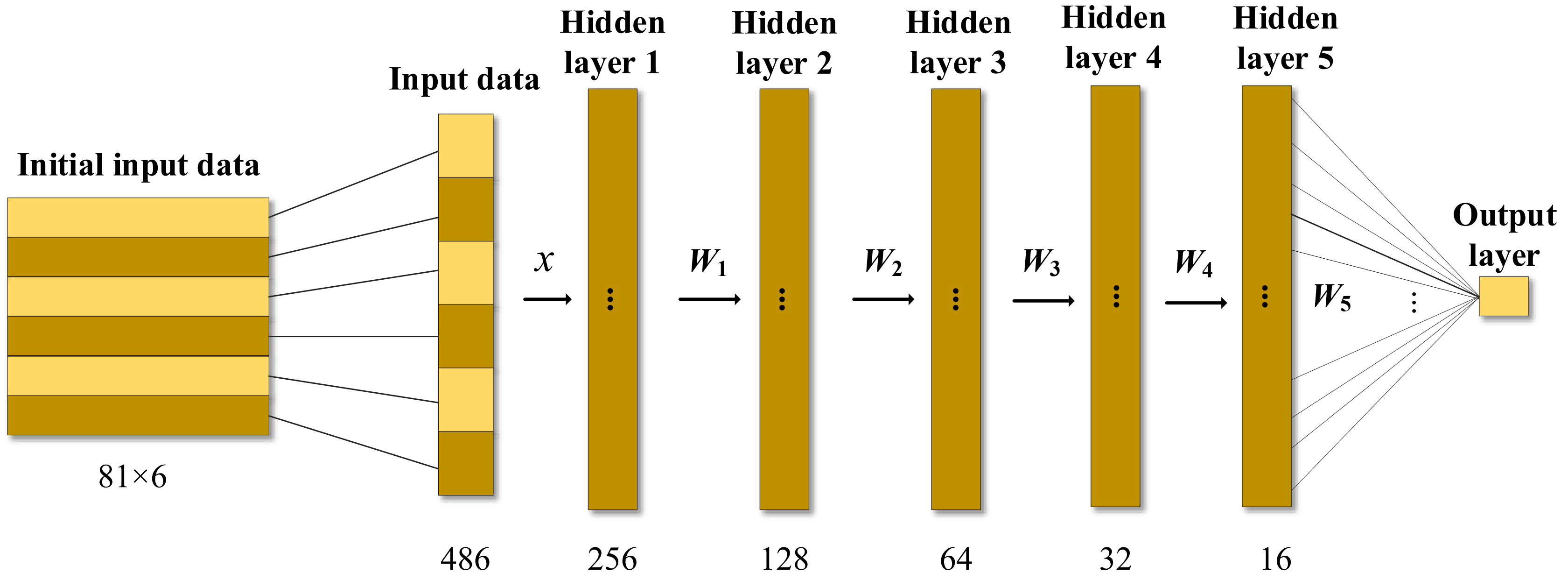
Generate the "target dataset" using EMT simulation tools
(consistent with the fault features and format of the field data)
to greatly augment the dataset for training

Step 3

Train the neural network using the "target dataset"
to achieve fault location compatible with field data

Step 3: Train the neural network

For the neural network (NN), the Basic Multilayer Perceptron (MLP) is applied as an example.



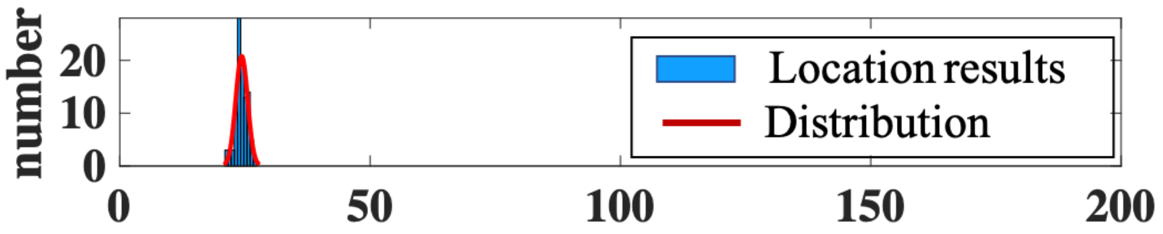
Batch size = 128
Epoch = 70
Learning rate = 0.01
Learning rate decay = 0.95
Loss function = L2 loss

Distribution map of fault location results

To avoid the randomness for each training (selection of training/validation datasets, batches, initial values, etc.)

→ 100 times of fault location results
→ Mean value

Example result



Outline

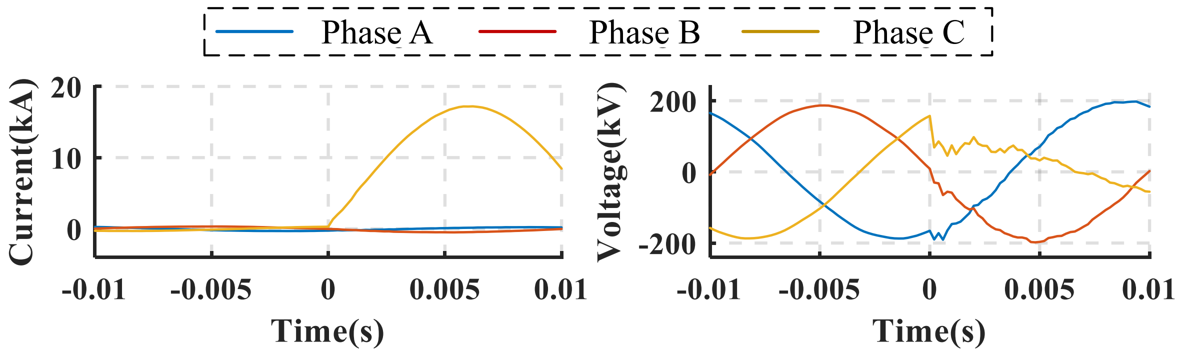
- Challenges of Data Driven Approaches for Fault Location
- Merging the Gap: Physics-Informed Data Driven Method
- Field Data Results
- Discussion
- Conclusion & Future Work

Field data results: Case 1

220 kV, 22.60 km Transmission Line
Fault Event: C-G fault, at 11.9 km of the Line

Field Data Stored in COMTRADE file:

- Measurements: 3 phase V & I at the local terminal
- Sampling Rate: 4 kHz
- Available data time window: **half a cycle** before and after the fault



Step 1: *Estimate system parameters*

(1) Equivalent source impedance

$Z_{s\alpha}$ = 1.2678+j6.7522 ohm, Z_{s0} = 1.5215+j11.3984 ohm

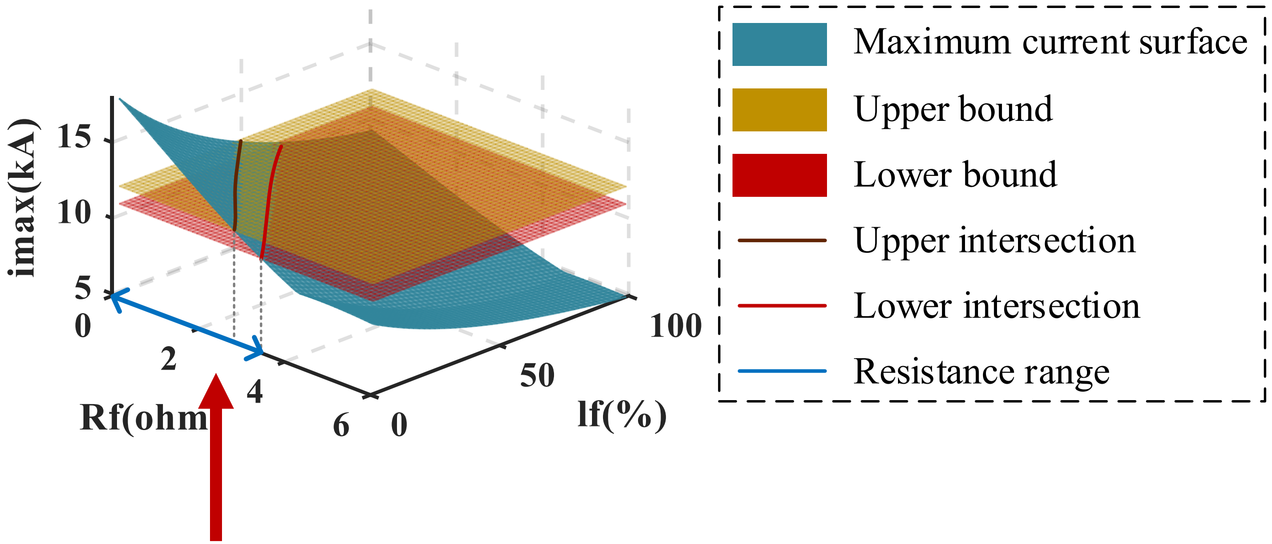
(2) Loading condition

Loading condition = 1.9916 deg

(3) Fault inception angle

FIA = 57.6 deg

(4) Fault resistance R_f (admittance Y_f)



R_f = 0-3.3 ohm

Field data results: Case 1

Step 2: "Target dataset" generation via simulation (Using Matlab Simulink)

Full Data group

Parameters	Value
Source impedance/ohm	Z_α=Z_β=0.1+j1, 0.4+j1.5, 1+j5 Z_0=0.2+j1.5, 0.8+j3, 2+j6
Loading condition/deg	±2, ±6, ±10, ±12
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10, 20, 50, 100, 150, 200
Fault location/km	1, 2, 3, …, 22 km
FIA/deg	0, 45, 90, …, 270, 315 (step: 45 deg)

Estimated system parameters

$Z_{s\alpha}$ = 1.2678+j6.7522
Z_{s0} = 1.5215+j11.3984 ohm

Loading condition = 1.9916 deg

R_f = 0-3.3 ohm

FIA = 57.6 deg

Target dataset (to Minimize the Gap between the filed data and the simulation)

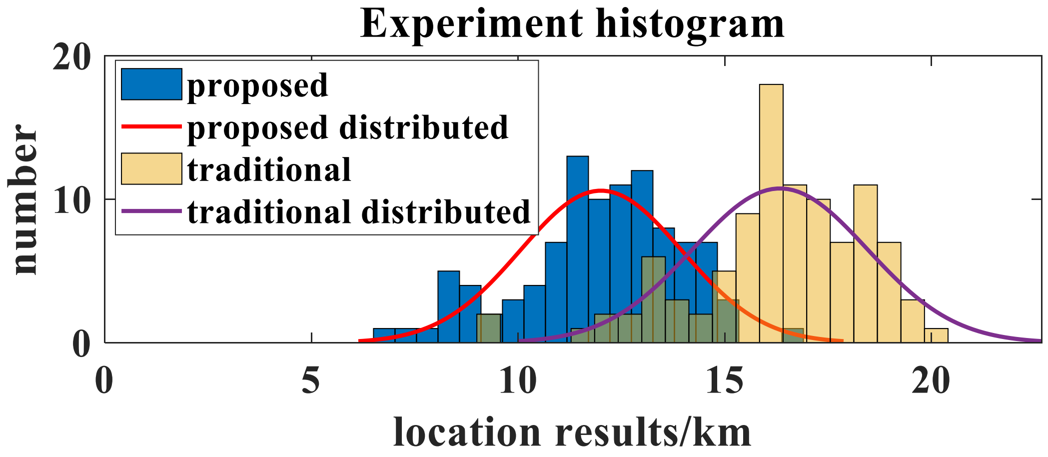
Parameters	Value
Source impedance/ohm	Z_α=Z_β= 0.4+j1.5, 1+j5; Z_0=0.8+j3, 2+j6
Loading condition/deg	±2
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10
Fault location/km	1, 2, 3, …, 22
FIA/deg	45, 90

Field data results: Case 1

Step** 3: **Train the neural network

→ 100 times of fault location results

→ Mean value



Fault Location Results (Actual value: 11.9 km)

Proposed data driven method:

12.003 km (error: **0.103 km**)

Traditional data driven method:

16.334 km (error: **4.434 km**)

Proposed Method presents much higher accuracy!

Implementation Platform:

Personal Computer, i7-7700 CPU

Implementation Software:

Parameter estimation: Matlab

Training procedure: Python

<u>Calculation Time:</u>

Parameter estimation: < 0.5 sec

Training and testing procedure for 1 time: < 1.2 sec

Overall time < 0.5 + 100 * 1.2 sec = 120.5 sec ≈ **2 min**

Calculation Burden is acceptable in practice!

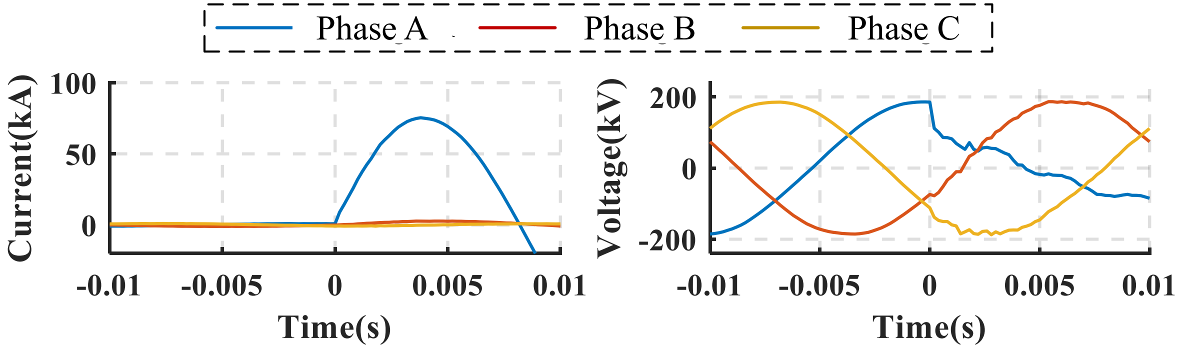
Field data results: Case 2

220 kV, 23.55 km Transmission Line

Fault Event: A-G fault, at 1.8 km of the Line

Field Data Stored in COMTRADE file:

- Measurements: 3 phase V & I at the local terminal
- Sampling Rate: 5 kHz
- Available data time window: **half a cycle** before and after the fault



Step 1: *Estimate system parameters*

(1) Equivalent source impedance

$Z_{s\alpha}$ = 0.2881+j0.8813 ohm, Z_{s0} = 0.1255+j1.0429 ohm

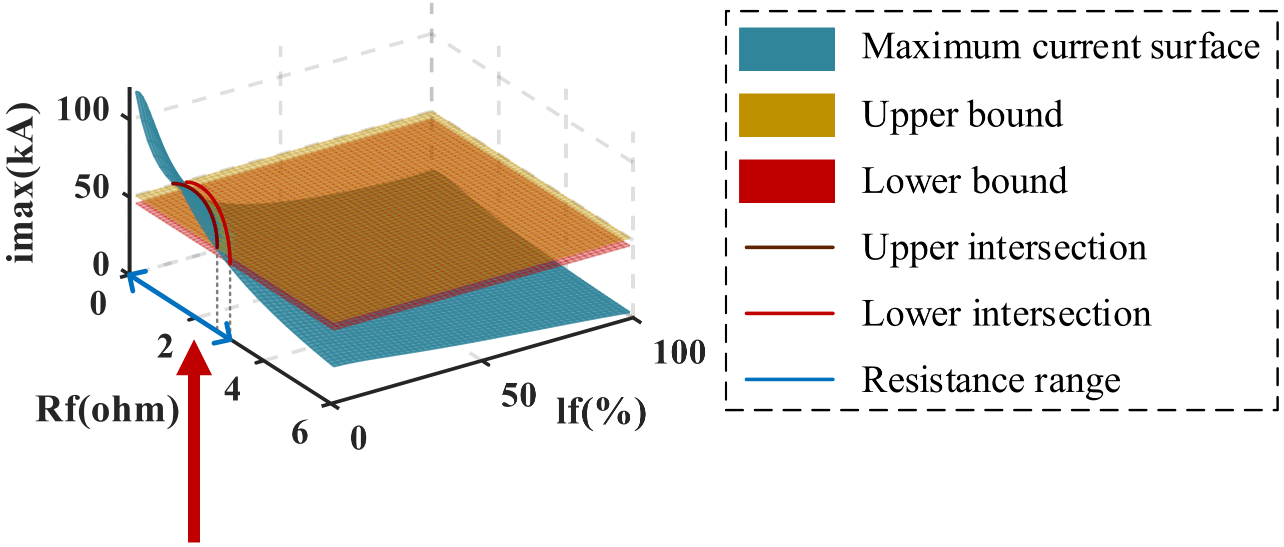
(2) Loading condition

Loading condition = 2.7417 deg

(3) Fault inception angle

FIA = 97.2 deg

(4) Fault resistance R_f (admittance Y_f)



R_f = 0-2.8 ohm

Field data results: Case 2

Step 2: "Target dataset" generation via simulation (Using Matlab Simulink)

Full Data group

Parameters	Value
Source impedance/ohm	$Z_\alpha=Z_\beta$=0.1+j1, 0.4+j1.5, 1+j5 Z_0=0.2+j1.5, 0.8+j3, 2+j6
Loading condition/deg	±2, ±6, ±10, ±12
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10, 20, 50, 100, 150, 200
Fault location/km	1, 2, 3, …, 23 km
FIA/deg	0, 45, 90, …, 270, 315 (step: 45 deg)

Estimated system parameters

$Z_{s\alpha}$ = 1.2678+j6.7522
Z_{s0} = 1.5215+j11.3984 ohm

Loading condition = 1.9916 deg

R_f = 0-3.3 ohm

FIA = 57.6 deg

Target dataset (to Minimize the Gap between the filed data and the simulation)

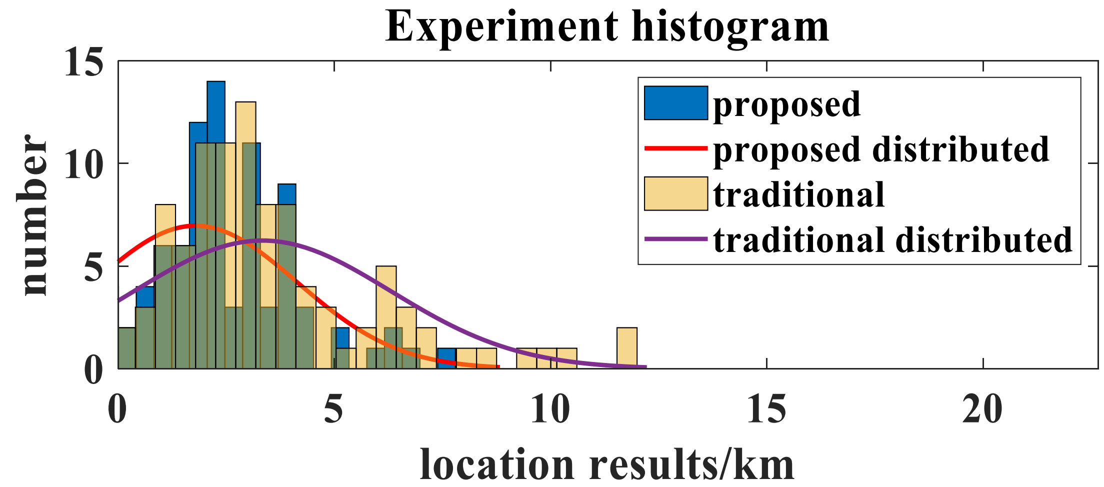
Parameters	Value
Source impedance/ohm	$Z_\alpha=Z_\beta$=0.1+j1, 0.4+j1.5; Z_0=0.2+j1.5, 0.8+j3
Loading condition/deg	2, 6
Fault resistance/ohm	0.01, 0.1, 0.2, 0.5, 1, 10
Fault location/km	1, 2, 3, …, 23
FIA/deg	45, 90

Field data results: Case 2

Step 3: Train the neural network

→ 100 times of fault location results

→ Mean value



Fault Location Results (Actual value: 1.8 km)

Proposed data driven method:

1.777 km (error: **0.023 km**)

Traditional data driven method:

3.337 km (error: **1.537 km**)

Proposed Method presents much higher accuracy!

Implementation Platform:

Personal Computer, i7-7700 CPU

Implementation Software:

Parameter estimation: Matlab

Training procedure: Python

Calculation Time:

Parameter estimation: < 0.5 sec

Training and testing procedure for 1 time: < 1.2 sec

Overall time < 0.5 + 100 * 1.2 sec = 120.5 sec ≈ **2 min**

Calculation Burden is acceptable in practice!

Outline

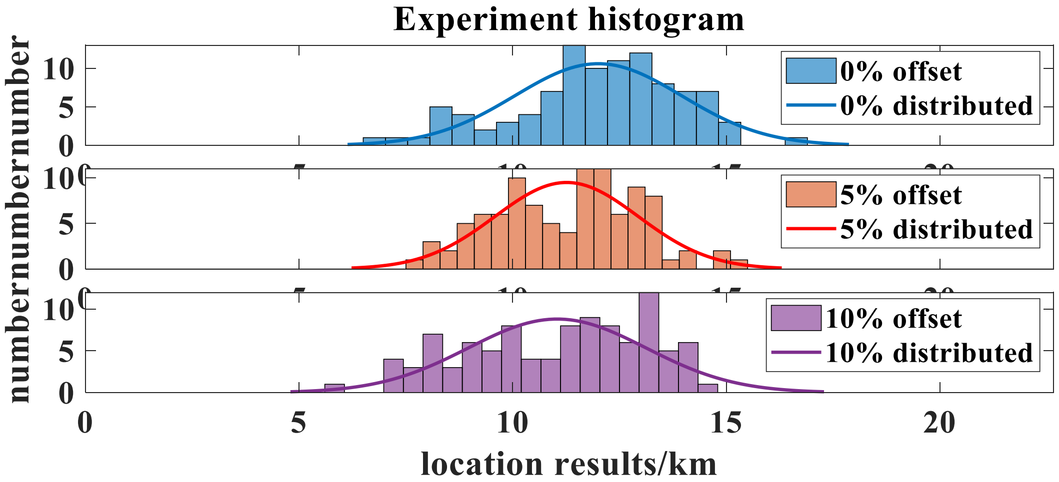
- ➢ Challenges of Data Driven Approaches for Fault Location
- ➢ Merging the Gap: Physics-Informed Data Driven Method
- ➢ Field Data Results
- ➢ Discussion
- ➢ Conclusion & Future Work

Discussion

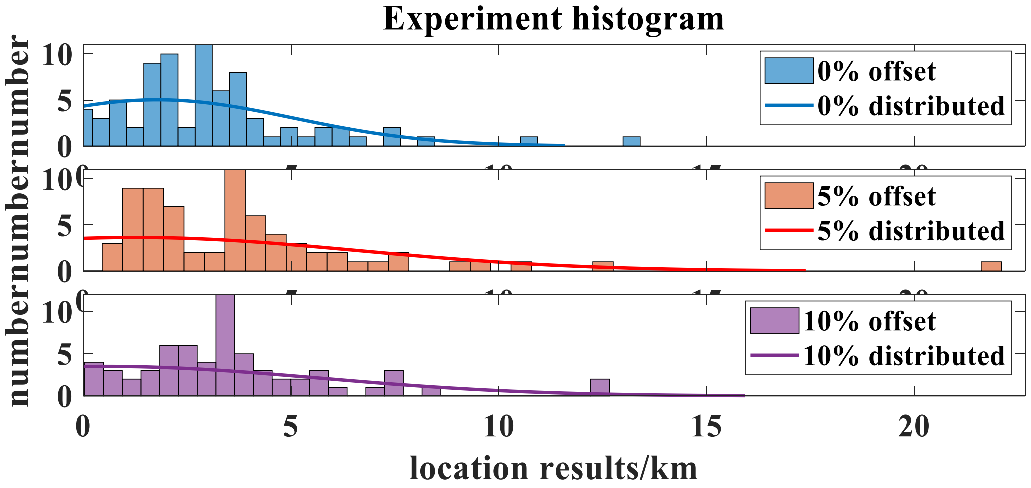
1. *Line Parameter Errors*

The **5%** **and 10% errors** are added to **all line parameters** in the utility database.

Field data results: Case 1



Field data results: Case 2



Fault Location Results (Actual value: 11.9 km)

- Proposed method with **0% parameter error**: 12.003 km (error: **0.103 km**)
- Proposed method with **5% parameter error**: 11.263 km (error: **0.637 km**)
- Proposed method with **10% parameter error**: 11.042 km (error: **0.858 km**)

Fault Location Results (Actual value: 1.8 km)

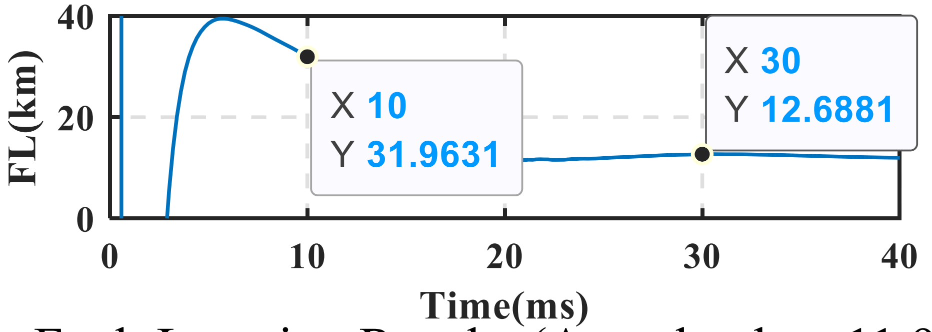
- Proposed method with **0% parameter error**: 1.777 km (error: **0.023 km**)
- Proposed method with **5% parameter error**: 1.248 km (error **0.552 km**)
- Proposed method with **10% parameter error**: 1.243 km (error: **0.557 km**)

Discussion

2. *Comparison to Existing Model based Single-Ended Method*

- Existing Takagi method (phasor based method)
- Phasors are extracted using IEEE C37.118 synchrophasor standard

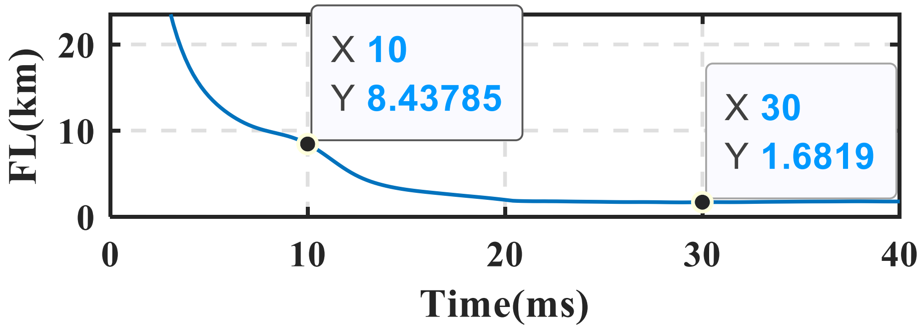
Field data results: Case 1



Fault Location Results (Actual value: 11.9 km)

- **Proposed** data driven method:
  12.003 km (error: **0.103 km**)
- Existing model based method, with **available time window of 0.5 cycle**:
  31.963 km (error: **20.063 km**)
- Existing model based method, with **available time window of 1.5 cycle**:
  12.688 km (error: **0.788 km**)

Field data results: Case 2



Fault Location Results (Actual value: 1.8 km)

- **Proposed** data driven method:
  1.777 km (error: **0.023 km**)
- Existing model based method, with **available time window of 0.5 cycle**:
  8.438 km (error: **6.638 km**)
- Existing model based method, with **available time window of 1.5 cycle**:
  1.682 km (error: **0.1181 km**)

Outline

- ➢ Challenges of Data Driven Approaches for Fault Location
- ➢ Merging the Gap: Physics-Informed Data Driven Method
- ➢ Field Data Results
- ➢ Discussion
- ➢ Conclusion & Future Work

Conclusion

- To apply data driven approach, one need to be **very careful** about **the dataset for training**. This is especially important for **fault related applications**.

- For **fault related applications**, the number of **field fault data** is **extremely limited**. We must use **simulation** to generate **data for training** (to ensure dataset completeness) and use **field data for testing** (to ensure practicability). Simulation data for both training and testing could present "too good but unrealistic" results.

- To generate proper **training dataset via simulation**, **field physics information** needs to be carefully taken into account. The **GAP** between **simulation** and **field data** could be **much reduced**.

- **If properly designed**, **physics-informed AI** could potentially be applied to **practical power systems**, and can **improve the performance** of traditional methods (higher accuracy, shorter data window, lower sampling rate).

Future Work

- **Minimize the assumption during the derivation** of parameters to improve estimation accuracy.

- Try to include **other information from the field** (remote side equivalent impedance, typical line loading conditions, if available) to enhance physics information.

- Find **more field fault data** for testing; try to include high resistance faults, faults with changing fault resistance, more fault types, etc.

Thank you!

Yu Liu, Associate Professor
ShanghaiTech University

Email: liuyu.shanghaitech@gmail.com
liuyu@shanghaitech.edu.cn

July 20, 2023