

Deep Learning based Digital Pre-Distortion for Power Amplifier Linearization

A Neural Network Approach to Nonlinear PA Compensation

Mrunali Bhamre

URSC,ISRO Internship

Project Report

December 2025

Contents

Abstract	3
1 Introduction	4
1.1 Background	4
1.2 Motivation	4
2 Technical Definitions	5
2.1 Digital Pre-Distortion (DPD)	5
2.2 Power Amplifier Nonlinearities	5
2.2.1 AM-AM Distortion	5
2.2.2 AM-PM Conversion	6
2.2.3 Memory Effects	6
2.3 Performance Metrics	6
2.3.1 Normalized Mean Square Error (NMSE)	6
2.3.2 Error Vector Magnitude (EVM)	7
2.3.3 Adjacent Channel Power Ratio (ACPR)	7
3 System Architecture	8
3.1 Overall Deep Learning Based DPD System	8
3.2 Physical Power Amplifier Model	8
3.3 Neural Network PA (NN-PA) Model	8
3.3.1 Input Feature Construction	9
3.3.2 NN-PA Architecture	9
3.4 Direct Learning of DPD Using NN-PA	9
3.4.1 DPD Feature Construction	10
3.4.2 Deep DPD Network Architecture	10
3.4.3 DPD Training Objective	10
3.4.4 Final Inference Chain	10
4 Performance Metrics	12
4.1 Normalized Mean Square Error (NMSE)	12

4.2	Error Vector Magnitude (EVM)	12
4.3	Adjacent Channel Power Ratio (ACPR)	13
4.4	AM-AM and AM-PM Characteristics	13
4.4.1	AM-AM Conversion	13
4.4.2	AM-PM Conversion	13
4.5	Gain and Linearity	13
5	Implementation Details	14
5.1	Baseband Signal Generation	14
5.2	Physical Power Amplifier Data Generation	14
5.3	Neural Network Power Amplifier Training	15
5.4	Fixed NN-PA Based DPD Evaluation Framework	15
5.5	DPD Model Training Using NN-PA	15
5.6	Training Hyperparameters	16
5.7	Inference and Evaluation	16
6	Experimental Results	17
6.1	Validation of NN-PA Model	17
6.2	DPD Performance Evaluation Framework	18
6.3	Comparative Analysis of DPD Models	18
7	Pathway to TWTA Implementation in Satellite Communications	20
7.1	TWTA Characteristics	20
7.1.1	Nonlinear Behavior	20
7.1.2	Operating Constraints	21
7.2	Step-by-Step Implementation Roadmap	21
7.2.1	Step 1: TWTA Characterization (1–2 months)	21
7.2.2	Step 2: Feedback Path Implementation (2–3 months)	21
7.2.3	Step 3: Enhanced Neural Network Design (2–3 months)	22
7.2.4	Step 4: Training Data Collection (1–2 months)	22
7.2.5	Step 5: Real-Time Processing Implementation (3–4 months)	22
7.2.6	Step 6: Adaptive Learning and Online Updates (2–3 months)	23
7.2.7	Step 7: Integration with Satellite Modem (2–3 months)	23
7.2.8	Step 8: Ground Station Validation (2–3 months)	23
8	Conclusion	24

Abstract

This report presents a deep learning-based power amplifier (PA) used for training digital pre-distortion (DPD) models to linearize power amplifiers in modern communication systems. The implementation employs a deep neural network to simulate the behaviour of a nonlinear physical PA exhibiting memory effects. Various DPD architectures, including feed forward neural networks, recurrent neural networks (RNN), gated recurrent units (GRU), and long short-term memory (LSTM) networks, are trained and compared to determine their suitability under identical PA conditions. Experimental results using 16-QAM signals demonstrate significant performance improvements, with NMSE reduction of 15–20 dB and EVM improvement of 20–25%. This study provides a systematic framework for comparing DPD models under controlled conditions and offers insights relevant to practical transmitter linearization in satellite and wireless communication systems.

Chapter 1

Introduction

1.1 Background

Power amplifiers are critical components in wireless and satellite transmitters. To maximize power efficiency, PAs are typically operated near saturation, where nonlinear behavior becomes significant. These nonlinearities lead to spectral regrowth, in-band distortion, and degradation in system capacity.

Digital Pre-Distortion (DPD) compensates for PA nonlinearities by applying an inverse characteristic at baseband prior to amplification.

1.2 Motivation

Traditional polynomial-based DPD techniques struggle to model strong nonlinearities and memory effects. Deep learning offers:

- Superior modeling of complex nonlinear behavior
- Data-driven adaptability
- Generalization across PA operating conditions
- Feasibility for real-time deployment

Chapter 2

Technical Definitions

2.1 Digital Pre-Distortion (DPD)

Digital Pre-Distortion (DPD) is a baseband signal processing technique that applies an inverse nonlinear transformation to the input signal such that the cascade of the pre-distorter and the power amplifier (PA) behaves approximately as a linear system.

If the PA is characterized by

$$y(n) = f_{\text{PA}}(x(n)), \quad (2.1)$$

then the DPD function $f_{\text{DPD}}(\cdot)$ satisfies

$$f_{\text{PA}}(f_{\text{DPD}}(s(n))) \approx G \cdot s(n), \quad (2.2)$$

where $s(n)$ is the desired signal and G is a constant gain.

2.2 Power Amplifier Nonlinearities

2.2.1 AM-AM Distortion

Amplitude-to-Amplitude (AM-AM) distortion describes the nonlinear relationship between the input and output signal amplitudes.

For an input signal

$$x(n) = A(n)e^{j\phi(n)}, \quad (2.3)$$

the output amplitude is given by

$$|y(n)| = g(A(n)), \quad (2.4)$$

where the nonlinear gain function is modeled as

$$g(A) = \frac{\alpha A}{1 + \beta A^2}. \quad (2.5)$$

2.2.2 AM-PM Conversion

Amplitude-to-Phase (AM-PM) conversion refers to phase distortion that depends on the signal amplitude. The output phase can be expressed as

$$\angle y(n) = \phi(n) + \psi(A(n)), \quad (2.6)$$

where the AM-PM distortion is modeled as

$$\psi(A) = -\gamma A^2. \quad (2.7)$$

2.2.3 Memory Effects

Memory effects occur when the PA output depends on both the current and past input samples. These effects arise from thermal dynamics, electrical memory, and trapping phenomena.

A PA with memory can be modeled as

$$y(n) = \sum_{m=0}^{M-1} h_m(|x(n-m)|) \cdot x(n-m). \quad (2.8)$$

2.3 Performance Metrics

2.3.1 Normalized Mean Square Error (NMSE)

Normalized Mean Square Error (NMSE) quantifies the normalized power of the error signal:

$$\text{NMSE} = \frac{\sum_{n=1}^N |y(n) - x(n)|^2}{\sum_{n=1}^N |x(n)|^2}. \quad (2.9)$$

In decibel scale:

$$\text{NMSE}_{\text{dB}} = 10 \log_{10}(\text{NMSE}). \quad (2.10)$$

2.3.2 Error Vector Magnitude (EVM)

Error Vector Magnitude (EVM) is defined as the RMS error between ideal and received constellation points:

$$\text{EVM}_{\text{RMS}} = \sqrt{\frac{\sum_{n=1}^N |S_n - R_n|^2}{\sum_{n=1}^N |S_n|^2}} \times 100\%. \quad (2.11)$$

2.3.3 Adjacent Channel Power Ratio (ACPR)

Adjacent Channel Power Ratio (ACPR) measures spectral regrowth:

$$\text{ACPR} = 10 \log_{10} \left(\frac{P_{\text{adjacent}}}{P_{\text{main}}} \right). \quad (2.12)$$

Chapter 3

System Architecture

3.1 Overall Deep Learning Based DPD System

The implemented system consists of three major processing blocks operating in cascade:

Desired Signal \rightarrow Deep DPD \rightarrow Neural Network PA \rightarrow Output Signal

3.2 Physical Power Amplifier Model

The reference PA is modeled as a nonlinear system with memory:

$$x_{\text{eff}}(n) = x(n) + \beta (0.5 x(n-1) + 0.2 x(n-2)) \quad (3.1)$$

The output amplitude and phase are distorted as:

$$g(n) = \frac{A_{\text{sat}} |x_{\text{eff}}(n)|}{1 + |x_{\text{eff}}(n)|^2} \quad (3.2)$$

$$\psi(n) = -\frac{\pi}{6} |x_{\text{eff}}(n)|^2 \quad (3.3)$$

The PA output is given by:

$$y_{\text{PA}}(n) = g(n) e^{j(\angle x_{\text{eff}}(n) + \psi(n))} \quad (3.4)$$

This model captures AM-AM compression, AM-PM conversion, and memory effects.

3.3 Neural Network PA (NN-PA) Model

To avoid repeated evaluation of the physical PA and to enable differentiable learning, a neural network is trained to emulate the PA behavior.

3.3.1 Input Feature Construction

For a memory depth M , the NN-PA input feature vector is defined as:

$$\mathbf{f}_{\text{PA}}(n) = [\Re\{x(n)\}, \dots, \Re\{x(n-M)\}, \Im\{x(n)\}, \dots, \Im\{x(n-M)\}] \quad (3.5)$$

The resulting input dimension is:

$$D_{\text{in}} = 2(M + 1) \quad (3.6)$$

3.3.2 NN-PA Architecture

The NN-PA is implemented as a fully connected feedforward neural network:

- Hidden layers: [128, 64, 32]
- Activation: LeakyReLU ($\alpha = 0.2$)
- Dropout rate: 0.2
- Separate output heads for real and imaginary components

The NN-PA mapping is:

$$\hat{y}(n) = \begin{bmatrix} \hat{I}(n) \\ \hat{Q}(n) \end{bmatrix} = \text{NN}_{\text{PA}}(\mathbf{f}_{\text{PA}}(n)) \quad (3.7)$$

The NN-PA is trained using mean squared error:

$$\mathcal{L}_{\text{PA}} = \frac{1}{N} \sum_{n=1}^N |\hat{y}(n) - y_{\text{PA}}(n)|^2 \quad (3.8)$$

3.4 Direct Learning of DPD Using NN-PA

In this work, the Digital Pre-Distortion (DPD) model is trained using a *direct supervised learning approach* based on a *Neural Network Power Amplifier (NN-PA)* surrogate model. The NN-PA is first trained to accurately emulate the nonlinear and memory behavior of the physical power amplifier. Once trained, the NN-PA replaces the physical PA during DPD training.

The learning objective is therefore:

$$f_{\text{NN-PA}}(f_{\text{DPD}}(x(n))) \approx x(n). \quad (3.9)$$

3.4.1 DPD Feature Construction

Let $x(n)$ denote the original complex baseband input signal. For a memory depth M , the DPD input feature vector is constructed as:

$$\mathbf{f}_{\text{DPD}}(n) = [\Re\{x(n)\}, \dots, \Re\{x(n-M)\}, \Im\{x(n)\}, \dots, \Im\{x(n-M)\}]. \quad (3.10)$$

The DPD network outputs a pre-distorted signal:

$$x_{\text{pre}}(n) = \hat{x}_{\text{I}}(n) + j\hat{x}_{\text{Q}}(n). \quad (3.11)$$

3.4.2 Deep DPD Network Architecture

The Deep DPD model is implemented as a memory-aware neural network consisting of:

- Input batch normalization
- Fully connected layers: [128, 64, 32]
- LeakyReLU activation functions
- Dropout regularization
- Separate output heads for in-phase (I) and quadrature (Q) components

The DPD mapping is expressed as:

$$x_{\text{pre}}(n) = f_{\text{DPD}}(\mathbf{f}_{\text{DPD}}(n)). \quad (3.12)$$

3.4.3 DPD Training Objective

The DPD network is trained by minimizing the error between the NN-PA output and the original input signal:

$$\mathcal{L}_{\text{DPD}} = \frac{1}{N} \sum_{n=1}^N |f_{\text{NN-PA}}(x_{\text{pre}}(n)) - x(n)|^2, \quad (3.13)$$

where

$$x_{\text{pre}}(n) = f_{\text{DPD}}(x(n)). \quad (3.14)$$

3.4.4 Final Inference Chain

During inference, the trained DPD is cascaded with the physical power amplifier:

$$x(n) \xrightarrow{\text{DPD}} x_{\text{pre}}(n) \xrightarrow{\text{Physical PA}} y_{\text{out}}(n) \approx G \cdot x(n). \quad (3.15)$$

During training and simulation, the NN-PA is used in place of the physical PA:

$$x(n) \xrightarrow{\text{DPD}} x_{\text{pre}}(n) \xrightarrow{\text{NN-PA}} \hat{y}(n). \quad (3.16)$$

Chapter 4

Performance Metrics

4.1 Normalized Mean Square Error (NMSE)

The Normalized Mean Square Error (NMSE) quantifies the modeling accuracy by measuring the normalized power of the error between the reference output of the physical PA and the modeled output.

Let $y(n)$ denote the reference PA output and $\hat{y}(n)$ denote the modeled or predicted output. The NMSE is defined as:

$$\text{NMSE} = 10 \log_{10} \left(\frac{\sum_{n=1}^N |y(n) - \hat{y}(n)|^2}{\sum_{n=1}^N |y(n)|^2} \right) \text{ dB}. \quad (4.1)$$

Lower NMSE values indicate better modeling accuracy.

4.2 Error Vector Magnitude (EVM)

Error Vector Magnitude (EVM) measures the distortion of a modulated signal by comparing the received constellation points with the ideal reference constellation.

The EVM is defined as:

$$\text{EVM}(\%) = \sqrt{\frac{\sum_{n=1}^N |y(n) - x(n)|^2}{\sum_{n=1}^N |x(n)|^2}} \times 100, \quad (4.2)$$

where $x(n)$ is the ideal transmitted signal and $y(n)$ is the received signal after amplification.

EVM is a critical metric for assessing compliance with wireless communication standards.

4.3 Adjacent Channel Power Ratio (ACPR)

Adjacent Channel Power Ratio (ACPR) evaluates the spectral regrowth caused by non-linear distortion in the PA.

ACPR is defined as the ratio of the power in the adjacent channel to the power in the main channel:

$$\text{ACPR} = 10 \log_{10} \left(\frac{P_{\text{adjacent}}}{P_{\text{main}}} \right) \text{ dB}. \quad (4.3)$$

Lower ACPR values indicate better spectral containment and reduced out-of-band emissions.

4.4 AM-AM and AM-PM Characteristics

4.4.1 AM-AM Conversion

AM-AM conversion characterizes the nonlinear relationship between the input signal amplitude and the output signal amplitude of the PA. It is defined as:

$$\text{AM-AM}(r) = |y(r)|, \quad (4.4)$$

where $r = |x(n)|$ is the input amplitude.

4.4.2 AM-PM Conversion

AM-PM conversion describes the phase distortion introduced by the PA as a function of input amplitude:

$$\text{AM-PM}(r) = \angle y(r) - \angle x(r). \quad (4.5)$$

These characteristics provide insight into the nonlinear behavior of the PA.

4.5 Gain and Linearity

The small-signal gain of the PA is defined as:

$$G = \frac{|y(n)|}{|x(n)|}. \quad (4.6)$$

An ideal linear PA exhibits constant gain over the entire operating range. Gain compression indicates nonlinear operation.

Chapter 5

Implementation Details

5.1 Baseband Signal Generation

A complex baseband signal is generated using 16-QAM modulation. Random symbols are mapped to constellation points, upsampled, and pulse-shaped using a Root Raised Cosine (RRC) filter.

The signal is peak-normalized to control the input back-off applied to the power amplifier:

$$x(n) = \frac{x(n)}{\max |x(n)|}. \quad (5.1)$$

The same baseband signal is used across all experiments to ensure consistent evaluation.

5.2 Physical Power Amplifier Data Generation

The physical PA is modeled using a nonlinear system with memory. Memory effects are introduced by combining delayed input samples:

$$x_{\text{eff}}(n) = x(n) + \alpha_1 x(n-1) + \alpha_2 x(n-2). \quad (5.2)$$

AM-AM and AM-PM nonlinearities are then applied to generate the PA output. The resulting input-output data is used exclusively for training the NN-PA model.

5.3 Neural Network Power Amplifier Training

A Neural Network Power Amplifier (NN-PA) is trained using supervised learning to emulate the behavior of the physical PA. The NN-PA learns the mapping:

$$y_{\text{PA}}(n) \approx f_{\text{NN-PA}}(x(n)). \quad (5.3)$$

The NN-PA is implemented as a feedforward neural network with memory features. Mean Squared Error (MSE) is used as the loss function:

$$\mathcal{L}_{\text{PA}} = \frac{1}{N} \sum_{n=1}^N |\hat{y}(n) - y(n)|^2. \quad (5.4)$$

Once trained, the NN-PA parameters are frozen and the model is used as a fixed surrogate PA for all subsequent DPD experiments.

5.4 Fixed NN-PA Based DPD Evaluation Framework

To ensure a fair comparison between different DPD architectures, the following evaluation strategy is adopted:

- The NN-PA model is trained once and kept fixed
- Identical baseband input signals are used for all DPD models
- Training hyperparameters are kept constant across DPD models
- Only the DPD architecture is varied

This framework isolates the effect of the DPD model architecture and ensures that performance differences are attributable solely to the DPD design.

5.5 DPD Model Training Using NN-PA

Each DPD model is trained directly as a pre-distorter using the fixed NN-PA. For an input signal $x(n)$, the DPD generates a pre-distorted signal $x_{\text{pre}}(n)$, which is passed through the NN-PA.

The training objective enforces linearization of the cascade:

$$\mathcal{L}_{\text{DPD}} = \frac{1}{N} \sum_{n=1}^N |f_{\text{NN-PA}}(x_{\text{pre}}(n)) - x(n)|^2. \quad (5.5)$$

The following DPD architectures are implemented and evaluated:

- Feedforward Neural Network (FFNN)
- Recurrent Neural Network (RNN)
- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)

All models use the same input feature construction and output mapping.

5.6 Training Hyperparameters

The training hyperparameters used for all NN-PA and DPD models are summarized in Table 5.1.

Table 5.1: Training Hyperparameters

Parameter	Value
Memory depth (M)	5
Hidden layers	[128, 64, 32]
Activation function	LeakyReLU
Dropout rate	0.2
Batch size	128
Training epochs	100
Optimizer	Adam
Learning rate	10^{-3}

5.7 Inference and Evaluation

During inference, the trained DPD is cascaded with the physical PA to evaluate real-world linearization performance. NMSE, EVM, and ACPR metrics are computed using identical test signals for all DPD models.

This ensures a consistent and unbiased comparison of DPD architectures.

Chapter 6

Experimental Results

This chapter presents the experimental results obtained for the power amplifier (PA) modeling and digital pre-distortion (DPD) framework. The results are organized to first validate the accuracy of the Neural Network Power Amplifier (NN-PA) model against the physical PA, followed by a comparative evaluation of multiple DPD architectures using a fixed NN-PA model.

Standard performance metrics such as NMSE, EVM, and ACPR are used for quantitative analysis.

6.1 Validation of NN-PA Model

Before evaluating DPD performance, the accuracy of the trained NN-PA is validated against the physical PA. This step is essential to ensure that the NN-PA can reliably act as a surrogate PA model during DPD training.

Table 6.1 compares the performance of the physical PA and the NN-PA using identical input signals and evaluation conditions.

Table 6.1: Performance Comparison Between Physical PA and NN-PA

Model	NMSE (dB)	EVM (%)	ACPR (dB)
Physical PA	-10.58	26.47	-43.01
NN-PA	-11.65	23.67	-37.37

The NN-PA achieves improved NMSE and EVM compared to the physical PA, demonstrating accurate modeling of nonlinear amplitude and phase distortions. Although a slight degradation in ACPR is observed, the results confirm that the NN-PA sufficiently captures PA behavior and can be used as a fixed surrogate model for DPD training.

6.2 DPD Performance Evaluation Framework

After NN-PA validation, the NN-PA model is frozen and used consistently across all DPD experiments. This ensures that the impact of different DPD architectures can be evaluated independently of PA modeling variations.

All DPD models are trained and tested using:

- The same baseband input signal
- Identical training and testing datasets
- Fixed NN-PA parameters
- Common training hyperparameters

This controlled setup guarantees a fair comparative analysis.

6.3 Comparative Analysis of DPD Models

Multiple DPD architectures are implemented and evaluated to assess their effectiveness in compensating PA nonlinearities. The following models are considered:

- Feedforward Neural Network (FFNN)
- Recurrent Neural Network (RNN)
- Long Short-Term Memory (LSTM)
- Gated Recurrent Unit (GRU)

Table 6.2 summarizes the performance of each DPD model in terms of NMSE, EVM, and ACPR.

Table 6.2: Comparative Performance of Different DPD Models

DPD Model	NMSE (dB)	EVM (%)	ACPR (dB)
No DPD	-10.58	26.47	-43.01
FFNN-DPD	-24.64	5.18	-35.88
RNN-DPD	-21.08	8.47	-31.79
LSTM-DPD	-29.17	2.89	-37.67
GRU-DPD	-30.78	2.39	-38.14

Without DPD, the PA showed severe distortion. Incorporating DPD improved all metrics:

Insights from DPD Performance Comparison

– **FFNN-DPD:**

- * Significantly improves NMSE and EVM compared to no DPD.
- * Limited ability to model memory effects results in moderate ACPR performance.

– **RNN-DPD:**

- * Captures some temporal dependencies, leading to better handling of PA memory than FFNN.
- * EVM and NMSE are slightly worse than FFNN, while ACPR is also lower, indicating trade-offs in in-band vs. out-of-band linearization.

– **LSTM-DPD (best performance):**

- * Achieves the lowest EVM and NMSE among all models, demonstrating strong capability to model both nonlinearities and memory effects.
- * ACPR improvement is also significant, showing balanced in-band and out-of-band linearization.

– **GRU-DPD (best trade-off):**

- * Slightly better NMSE and EVM than LSTM, with comparable ACPR.
- * Offers reduced computational complexity and training time, making it efficient for practical implementation.

– **General Observation:**

- * Recurrent architectures (LSTM, GRU) outperform feedforward models due to their ability to capture temporal dependencies.
- * There is a trade-off between in-band distortion (EVM, NMSE) and out-of-band emissions (ACPR); model choice can balance performance vs. complexity.

Chapter 7

Pathway to TWTA Implementation in Satellite Communications

7.1 TWTA Characteristics

Traveling Wave Tube Amplifiers (TWTAs) are the dominant high-power amplification devices used in satellite communication systems due to their high efficiency and ability to operate at microwave and millimeter-wave frequencies. However, TWTAs exhibit strong nonlinear behavior and long memory effects, making linearization particularly challenging.

Key characteristics of TWTAs include:

- Severe AM-AM compression compared to solid-state PAs
- Strong AM-PM conversion with phase distortion up to 30° – 40°
- Long thermal memory effects with time constants ranging from milliseconds to seconds
- Hysteresis effects depending on envelope dynamics

7.1.1 Nonlinear Behavior

TWTA nonlinearities are more pronounced than those observed in solid-state amplifiers. The AM-AM characteristic exhibits early saturation, while AM-PM distortion introduces significant phase rotation even at moderate input power levels. Additionally, memory effects arise due to thermal variations, charge trapping, and bias network dynamics.

These effects result in:

- Increased spectral regrowth
- Constellation rotation and warping
- Reduced adjacent channel isolation

7.1.2 Operating Constraints

Practical TWTA deployment in satellites imposes strict operational constraints:

- High DC power consumption (100 W to several kW)
- Limited dynamic range and output back-off
- Temperature-dependent characteristics due to orbital thermal cycling
- Aging and performance drift over satellite lifetimes exceeding 15 years

These constraints necessitate robust, adaptive, and low-complexity DPD solutions.

7.2 Step-by-Step Implementation Roadmap

7.2.1 Step 1: TWTA Characterization (1–2 months)

The first step involves comprehensive characterization of the TWTA:

- Measurement of static AM-AM and AM-PM curves
- Identification of effective memory depth
- Temperature-dependent behavior analysis
- Definition of safe operating regions

Typical test signals include single-tone, two-tone, and modulated carriers such as QPSK, 8PSK, and 16-APSK.

7.2.2 Step 2: Feedback Path Implementation (2–3 months)

DPD training requires a feedback path to observe the TWTA output. Required components include:

- Wideband observation receiver (bandwidth $\geq 2 \times$ signal bandwidth)
- High-resolution ADCs (14–16 bits)

- Time and phase alignment mechanisms

In satellite systems where direct feedback is unavailable, alternative solutions include pre-launch characterization, telemetry-based adaptation, and indirect estimation from link performance metrics.

7.2.3 Step 3: Enhanced Neural Network Design (2–3 months)

To address long memory effects in TWTAs, the DPD neural network must be enhanced:

- Increase memory depth from 5 to 20–50 taps
- Employ recurrent architectures such as LSTM or GRU
- Use multi-timescale models combining fast and slow memory branches

The enhanced DPD mapping can be expressed as:

$$\hat{x}(n) = \text{DNN} \left(\{y(n-k)\}_{k=0}^{M_1}, \{|y(n-k)|\}_{k=0}^{M_2}, T_{\text{TWTA}}(n) \right) \quad (7.1)$$

7.2.4 Step 4: Training Data Collection (1–2 months)

Robust DPD training requires extensive datasets:

- 10–100 million samples across operating conditions
- Multiple modulation formats (QPSK, 8PSK, 16/32-APSK)
- Various input back-off levels
- Temperature and aging variations

Synthetic data generated using physics-based TWTA models can supplement real measurements.

7.2.5 Step 5: Real-Time Processing Implementation (3–4 months)

Candidate hardware platforms include:

- **FPGA**: Low latency, fixed-point implementation, moderate power consumption
- **GPU/TPU**: Suitable for ground stations and gateway processing

- **ASIC**: Radiation-hardened, ultra-low power for space deployment

Latency constraints typically require DPD processing within 100 ns.

7.2.6 Step 6: Adaptive Learning and Online Updates (2–3 months)

Long-term satellite operation demands adaptive DPD:

- Transfer learning with frozen feature layers
- Meta-learning for rapid adaptation
- Online gradient descent with very small learning rates

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} L(y_t, s_t), \quad \eta \ll 1 \quad (7.2)$$

7.2.7 Step 7: Integration with Satellite Modem (2–3 months)

The DPD block is integrated between the modem and RF upconverter while preserving:

- Frame timing and synchronization
- Variable coding and modulation (VCM)
- Automatic gain control (AGC) stability

7.2.8 Step 8: Ground Station Validation (2–3 months)

Final validation includes:

- Hardware-in-the-loop testing
- Spectral mask compliance verification
- End-to-end BER and link budget evaluation
- Live traffic testing with monitoring

Chapter 8

Conclusion

This work presents a systematic investigation of deep learning-based digital pre-distortion (DPD) for power amplifier linearization using a neural network-based power amplifier (NN-PA) as a surrogate for the physical PA.

The NN-PA accurately emulated the physical PA with only 1.07 dB NMSE deviation and an EVM improvement of nearly 3%, validating its use for controlled DPD evaluation.