

2023 年度卒業

修士論文

深層学習による動画予測手法を用いた太陽全球紫外線像
の時系列予測

Time-Series Prediction of Full-Disk Solar Ultraviolet Images
Using Deep Learning-based Video Prediction Method

所属	新潟大学 大学院自然科学研究科 電気情報工学専攻 情報工学コース 飯田研究室
氏名	佐々木明良
学籍番号	F22C017D

概要

アブストラクト

目次

第 1 章 研究背景	4
第 2 章 動画予測	6
2.1 動画予測の定義と定式化	6
2.1.1 動画予測モデル	6
2.1.2 最適化	7
2.2 動画予測のための基礎技術	7
2.2.1 Convolutional Neural Network (CNN)	7
2.2.1.1 疊み込み	8
2.2.1.2 プーリング	8
2.2.2 Encoder-Decoder	9
2.2.2.1 エンコーダ	10
2.2.2.2 デコーダ	10
2.2.3 Recurrent Neural Network (RNN)	10
2.2.4 Long Short Term Memory (LSTM)	11
2.2.5 Attention	12
2.3 動画予測フレームワーク	13
2.3.1 ConvLSTM	14
2.3.2 PredRNN	14
2.3.2.1 時空間メモリフロー	15
2.3.2.2 時空間 LSTM ユニット (ST-LSTM)	16
2.3.3 Motion-Aware Unit(MAU)	17
2.3.3.1 アーキテクチャ	18
2.3.3.2 主な実験結果	20
第 3 章 データ	22
3.1 SDO / AIA	22
3.1.1 AIA 211 Å	22

3.1.2	AIA 193 Å	23
3.1.3	AIA 171 Å	24
3.2	前処理	25
3.2.1	不正な画像の除去	26
3.2.2	スケーリングと正規化	26
3.2.3	データセットの分割	29
第 4 章	Motion-Aware Unit を用いた 1 波長を入力とした紫外線像の全球時系列予測	31
4.1	実験概要	31
4.2	実験設定	31
4.3	学習の推移	32
4.4	実験結果	33
4.4.1	全球での評価	35
4.4.1.1	平均輝度の再現	36
4.4.1.2	画像類似度	40
4.4.2	経度依存性の評価	42
4.4.2.1	平均輝度の再現	42
4.4.2.2	画像類似度	44
4.4.3	東側外縁部に対する評価	46
4.4.3.1	活動領域に対する視覚的評価	46
4.4.3.2	予測対実測散布図による定量的評価	49
4.5	考察	50
4.5.1	全球での性能	51
4.5.2	経度に対するロバスト性	51
4.5.3	東側外縁部における予測性能	52
4.5.4	複雑なシステムに対する確率的予測の有効性	53
第 5 章	Motion-Aware Unit を用いた 3 波長を入力とした紫外線像の全球時系列予測	56
5.1	実験概要	56
5.2	実験設定	56
5.3	学習の推移	58
5.4	実験結果	58
5.4.1	全球での評価	58
5.4.1.1	平均輝度の再現	59
5.4.1.2	画像類似度	59
5.4.2	経度依存性の評価	61
5.4.2.1	平均輝度の再現	61

5.4.2.2	画像類似度	63
5.4.3	東側外縁部に対する評価	63
5.4.3.1	視覚的評価	63
5.4.3.2	散布図	63
5.4.4	まとめ	66
5.5	考察	66
第 6 章	まとめ	69

第1章

研究背景

宇宙天気とは、太陽活動に起因する宇宙環境の現象を指し、激しい宇宙天気の変動は地球上の電力網や衛星通信など、人間の技術システムに影響を及ぼすことがある。太陽の表面や大気における物理現象、特に太陽フレアやコロナ質量放出（CME）などの爆発的なイベントは、地球に到達する高エネルギー粒子や放射線の量を増加させ、宇宙天気の変動に大きな影響を与えることが知られている。そのため、太陽活動を観測し、宇宙天気を予測することは、人間の技術システムを宇宙天気の影響から守るための重要な課題である。

宇宙天気の予報には、太陽フレアの爆発や CME の発生を予測するものや、太陽風の地球到達時刻や強度を予測するものなどがある。これらの予測は、様々な観測機器によって得られる太陽観測データを用いて行われるが、紫外線像はその中でも重要な情報源の一つである。（DeFN、WindNet の話とか）このように、太陽の紫外線像、およびそれから得られる特徴量は、宇宙天気予報において重要な役割を果たしている。そこで、まだ観測されていない未来の紫外線画像を予測、生成することで、より早期の宇宙天気予報の実現に貢献できるのではないかと考えた。

近年、深層学習技術の発展により、「動画予測（Video Prediction）」と呼ばれる技術が注目されている。動画予測とは、動画の一部を入力として、それに続く未来の動画を予測するタスクである。動画予測を行う深層学習モデルのアーキテクチャの多くは、Long Short-Term Memory(LSTM) などの再帰的ニューラルネットワーク (RNN) のアーキテクチャを基本とする。さらに、特徴量抽出および伝播に Convolutional Neural Network(CNN) を用いることで空間的な特徴を時系列にわたってとらえ、Decoder モデルによって動画を生成する。このような動画予測モデルは、Conv-LSTM の登場で初めて提案されて以降、様々なモデルが提案されており、自動運転や天気予報など、様々な分野での応用が期待されている。

本研究では、動画予測モデルを用いて、数日後の紫外線画像を予測、生成することを目的とする。Deep Flare Net などの多くの深層学習を用いた宇宙天気予報モデル、またそれらを用いた人による主観的な宇宙天気予報は、現在の観測情報を用いて、数時間後から数日後までの宇宙天気を予

測するものが多い。それらの情報源として、数日後の高精度な紫外線画像を生成することができれば、より早期の宇宙天気予報の実現に貢献できるのではないかと考えた。

そのような数日後の紫外線像の生成のために、本研究では Motion-Aware Unit(MAU) と呼ばれる動画予測モデルを用いる。MAU は、RNN や CNN を用いた基本的な動画予測モデルを基本としつつ、各時間時点における画像の処理に MAU Cell と呼ばれるモジュールを多層的に積み重ねたアーキテクチャを採用している。MAU Cell は Attention と呼ばれる機構を持ち、長期的な依存関係を適切に学習する能力を持つ。また、従来の動画予測モデルで提案してきた Encoder-Decoder モデルや、メモリフローの改善など、様々な改良が加えられており、多くの動画予測モデルの中でも要求する計算量に対して高い予測精度を達成している。

本研究では、MAU を用いて、数日後の紫外線画像を予測、生成することを目的とするが、その評価には主に輝度強度の再現性を用いる。これは、宇宙天気予報モデルの多くではその特徴量として輝度強度を用いていること由来する。

第 2 章

動画予測

2.1 動画予測の定義と定式化

動画予測は、既知のビデオフレームの系列から未来のフレームを予測するタスクであり、教師なし学習、または自己教師あり学習の一種として位置付けられる。このタスクは、時空間的な連続性と一貫性を持つ未来のフレームシーケンスを生成することを目指す。

2.1.1 動画予測モデル

動画予測の目的は、与えられた過去のフレームシーケンスから未来のフレームを予測するモデル M を最適化することである。ほとんどの動画予測モデルにおいて、モデルの扱うシークエンスは入力長と出力長の二つに分割される。入力長は、モデルが予測を行うために必要な過去のフレームであり、一貫してモデルがその動画の時空間的ダイナミクスを学習するために使用される。出力長は、モデルが予測を行う未来のフレームであり、モデルが最終的に生成するフレームシーケンスの一部となり、この部分に対して損失が計算される。また、もっとも初めの出力フレーム以降のフレームでは、直前の出力フレームが入力フレームとして扱われる。動画予測モデル M は、出力長として位置付けられるある時間ステップ t のフレーム \hat{X}_t を生成する際は、その直前までの過去のフレーム $X_{0:t-1}$ を入力として受け取り、それに基づいてそれに続く未来のフレームを生成する。このプロセスは以下のように定式化される：

$$\hat{X}_t = M(X_{0:t-1}) \quad (2.1)$$

2.1.2 最適化

最適化プロセスは、一連の学習データセットを用いて、損失関数 L を最小化するように動画予測モデル M のパラメータを調整する。このプロセスは、以下のように表現される：

$$M_{\text{optimized}} = \operatorname{argmin}_M L(\hat{X}, X) \quad (2.2)$$

ここで、 $M_{\text{optimized}}$ は最適化された予測モデルを表す。この損失関数は、予測された未来のフレームと実際の未来のフレームとの差異を測定するために使用され、一般的には平均二乗誤差（Mean Squared Error, MSE）が用いられる。これは、予測されたフレームと実際のフレームのピクセル単位の差異を測定する。MSE は次のように定義される：

$$MSE = \frac{1}{N} \sum_{i=1}^N (X_i - \hat{X}_i)^2 \quad (2.3)$$

2.2 動画予測のための基礎技術

動画予測には、複数のフレーム間での時空間的な関連を捉え、未来のフレームを予測する能力が必要である。この目的を達成するためには、畳み込みニューラルネットワーク (CNN)、エンコーダ・デコーダ構造、長短期記憶 (LSTM)、そしてアテンションメカニズムといった複数の技術が組み合わされる。ここではそういう基礎技術群について説明する。

2.2.1 Convolutional Neural Network (CNN)

CNN は動画予測において、各フレームの空間的な特徴を抽出する役割を担う。動画は時間的な次元を持つ一連の画像であり、各フレームにおける空間的な特徴を理解することは、時間的な次元を解析する前の重要なステップである。

CNN は、Fukushima (1980) によるネオコグニトロンの開発や、LeCun et al. (1989) による手書き数字認識の研究などにより、画像認識の分野で初めて使用された。その後、Krizhevsky et al. (2012) により開発された AlexNet が成功したことにより、現在では画像認識の分野において最も広く用いられる深層学習の一形態となっている。CNN は、画像の局所的な特徴を捉るために、畳み込み層とプーリング層を交互に繰り返すことで構成される。畳み込み層は画像から特徴を抽出するためのフィルターの役割を果たし、プーリング層は画像の特徴を圧縮する役割を持つ。CNN を用いた多くの画像処理アプリケーションにおいては、この畳み込み操作とプーリング操作が連続的に繰り返されることで、入力データの高次元な特徴を抽出することを可能にしている。ここでは、その二つの重要な操作について説明する。

2.2. 動画予測のための基礎技術

2.2.1.1 署み込み

CNN の主な特徴は、局所的な特徴を効率的に捉えることができる署み込み層にある。この署み込み層は、入力された画像から特定の特徴を抽出するフィルターの役割を果たし、これにより画像の特徴を圧縮して表現することができる。

CNN における署み込み処理は、入力データに対する特定のカーネルの適用として理解される。カーネルは小規模な行列であり、入力データの局所的な領域に対して適用される。この署み込み操作は、入力データを I 、カーネルを K とした場合、以下の式で表される。

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n)K(m, n) \quad (2.4)$$

この操作の様子を図 2.1 に示す。この操作は、入力データの全域にわたって繰り返され、最終的に特徴マップと呼ばれる新しい行列が生成される。特徴マップには、その入力データにおける特定のパターンや構造が抽出されている。

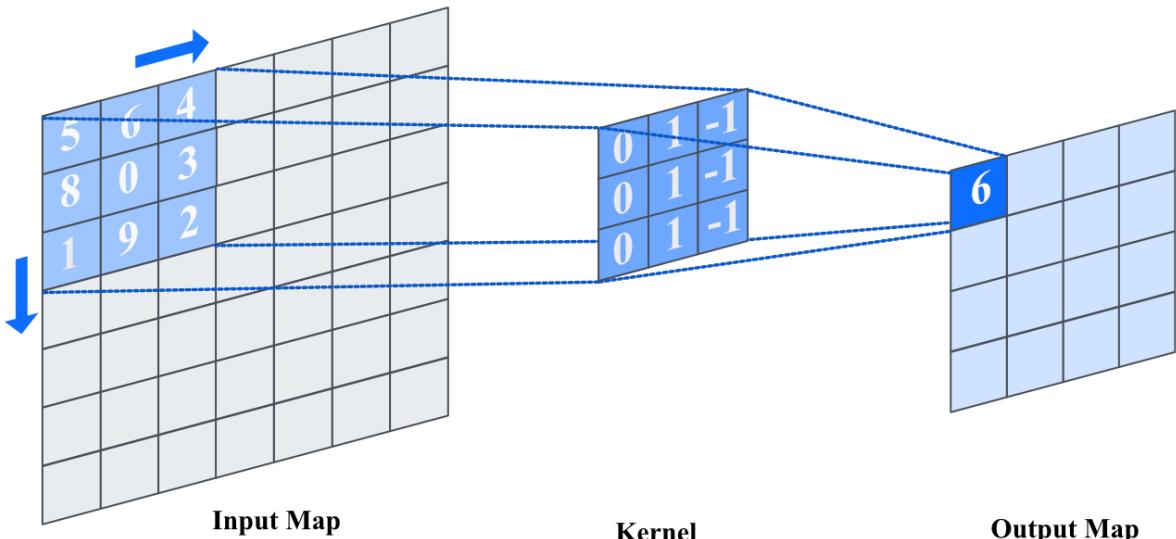


図 2.1: 署み込み操作の様子。左: 入力データ。中央: カーネル。右: 署み込み操作により生成された特徴マップ。

2.2.1.2 プーリング

この層の主な目的は、特徴マップの次元を減少させることである。具体的には、プーリング層は特徴マップの小さな領域を集約し、その領域内の代表的な値（最大値や平均値）を抽出する。この操

2.2. 動画予測のための基礎技術

作により、ネットワークは画像の局所的な変化に対してより頑健になり、より抽象的な特徴表現を学習することが可能である。

- **最大値プーリング (Max Pooling)** : この手法では、各領域の最大値が選択される。これにより、特徴マップから最も強い信号を保持し、関連性の低い信号を破棄する。最大値プーリングは、特に画像内のテクスチャや形状などの顕著な特徴を強調するのに有効である。ここで、入力となる特徴マップを F , プーリング領域のサイズを $m*n$, プーリング層の出力を P とすると、最大値プーリングは以下の式で表される。

$$P(i, j) = \max_m \max_n F(i + m, j + n) \quad (2.5)$$

最大値プーリングの様子を図 2.2 に示す。

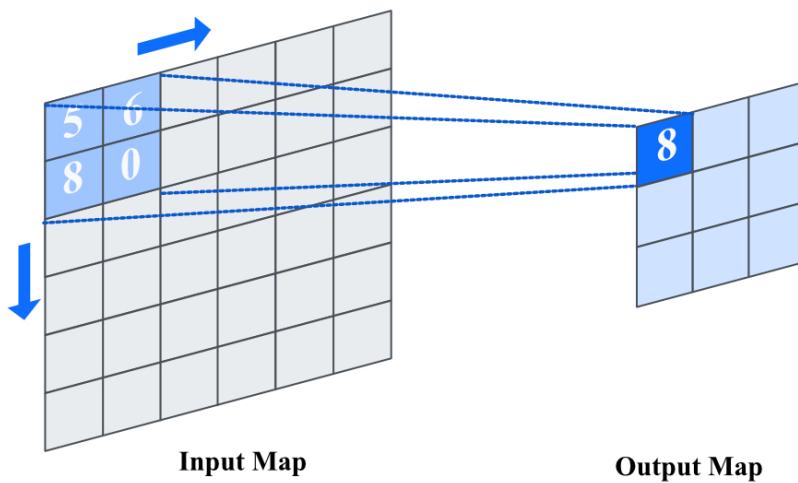


図 2.2: 最大値プーリングの様子。左: 入力データ。右: 最大値プーリングにより生成された特徴マップ。

- **平均値プーリング (Average Pooling)** : 平均プーリングは、各領域の平均値を計算する。これにより、特徴マップの全体的な特性をより平滑化し、より均一な特徴表現を提供することが可能である。ここで、入力となる特徴マップを F , プーリング層の出力を P とすると、平均値プーリングは以下の式で表される。

$$P(i, j) = \frac{1}{mn} \sum_m \sum_n F(i + m, j + n) \quad (2.6)$$

2.2.2 Encoder-Decoder

エンコーダ・デコーダ構造は画像処理において広く用いられ、特に U-Net のようなアーキテクチャが代表的である。エンコーダ・デコーダ構造は、一連の入力データを処理し、それを内部表現に変

2.2. 動画予測のための基礎技術

換するエンコーダ部分と、この内部表現から出力を生成するデコーダ部分の二つの主要なコンポーネントから構成される。動画予測において入力データとなるのは、動画の各フレームであり、出力データはそれに続く未来のフレームである。ここでは画像処理におけるエンコーダ・デコーダ構造に焦点を当てて説明する。

2.2.2.1 エンコーダ

エンコーダは入力データを CNN によって処理し、それを高次元から低次元の表現に変換する。このプロセスは、入力データに含まれる重要な情報を抽出し、より扱いやすいサイズまたは形式に圧縮することを目的とする。動画予測アプリケーションにおいては、空間的特徴に加え、複雑な時間的特徴の依存性をモデリングするため、一般的な画像処理ディープラーニングモデルと比較して大量の計算資源を必要とする。そのため、エンコーダにより特徴を圧縮し、より低い次元で高度な特徴を抽出することは、計算コストの削減という点からも非常に有用である。

2.2.2.2 デコーダ

デコーダは本質的にエンコーダの逆処理である。動画予測においては、その直前のアーキテクチャにより生成された内部表現を受け取り、目的とする出力を生成する役割を持つ。動画予測の再帰ネットワーク内では、エンコーダにより圧縮された行列が扱われるため、そのままでは出力には不適切である。デコーダはそのような圧縮された表現を元の次元に展開し、出力に適した形式に変換する。

2.2.3 Recurrent Neural Network (RNN)

リカレントニューラルネットワーク (RNN) は、時系列データや自然言語などのシーケンシャルな情報を扱うために開発されたニューラルネットワークの一種である (Werbos (1990))。RNN の特徴は、過去の情報を隠れ状態として保持し、それを利用して次の出力を生成する点にある。このような再帰的な構造を持つことにより、RNN は時系列性をもつデータを扱うことが可能である。RNN の構造を図 2.3 に示す。

RNN の隠れ層 h_t は、時刻 t での入力 x_t と、 $t - 1$ での隠れ層 h_{t-1} から、以下の式で表される。

$$h_t = f(W_{xh} \cdot x_t + W_{hh} \cdot h_{t-1} + b_h) \quad (2.7)$$

ここで、 W_{xh} と W_{hh} はそれぞれ入力と隠れ層の重み、 b_h はバイアスを表し、 f は活性化関数を表す。活性化関数には、通常、 \tanh や $ReLU$ などの非線形関数が用いられる。

このように、RNN は過去の情報を保持することで、時系列データの予測を行うことが可能である

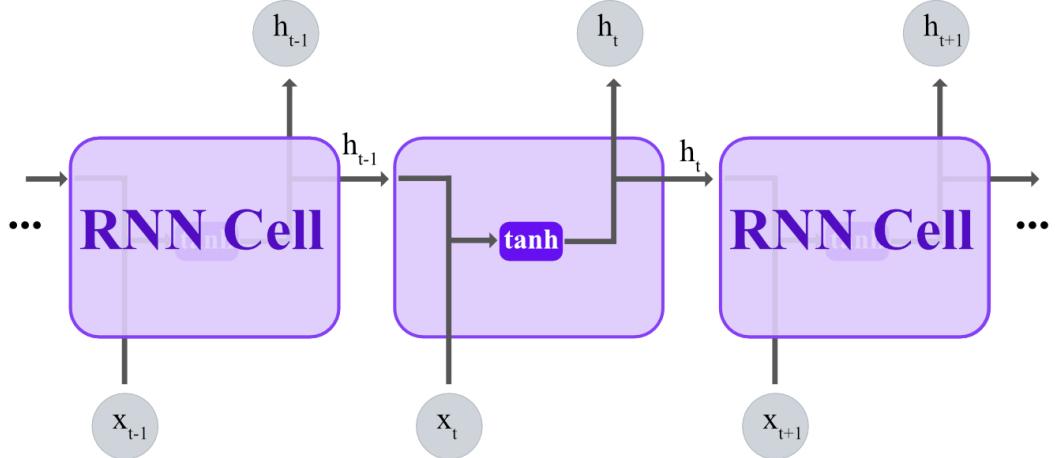


図 2.3: RNN の構造。緑色のブロック h は隠れ層を表す。

が、実際には、長期的な依存関係を捉えることに困難を抱えていた。この問題は勾配消失問題として知られる。これは、RNN の活性化関数の出力が各ステップで乗算されることにより、ネットワークが深くなるほど、勾配が指数関数的に消失してしまうことに起因する。

2.2.4 Long Short Term Memory (LSTM)

この勾配消失問題を解決するために Hochreiter and Schmidhuber (1997) によって開発されたのが、長短期記憶 (Long Short Term Memory, LSTM) である。LSTM の主要な特徴は、直前のセルの出力 h_t を次のセルで使用しながら、セル状態 c_t と呼ばれる長期的な情報を変化させることで、より長期的な情報を扱うことが可能である点にある。

LSTM の構造は以下の三つのゲートから成る：

1. 忘却ゲート (Forget Gate) : このゲートは、セル状態に含まれる情報の一部を削除する役割を担う。LSTM ネットワークが長期的な依存関係を学習する過程で、関連性の低い古い情報を捨てることが重要である。忘却ゲートはシグモイド関数を使用して、どの情報を保持し、どの情報を忘れるかを決定する。
2. 入力ゲート (Input Gate) : 入力ゲートは、新しい情報をどの程度セル状態に追加するかを決定する。このゲートでは、シグモイド関数がどの情報を更新するかを決定し、tanh 関数が新しい候補値を生成する。そして、これら二つの値の積が新しい情報としてセル状態に追加される。

2.2. 動画予測のための基礎技術

3. 出力ゲート (Output Gate) : 出力ゲートは、現在のセル状態に基づいて、ネットワークの出力を決定する。このゲートは、シグモイド関数を使用して、セル状態のどの部分が output されるべきかを決定し、 \tanh 関数によって処理されたセル状態との積が最終的な出力となる。

LSTM ユニットの数学的な定式化は以下の通りである。 h_t は時刻 t での隠れ状態、 c_t はセル状態、 x_t は入力、 f_t 、 i_t 、 o_t はそれぞれ忘却ゲート、入力ゲート、出力ゲートの活性化関数を表す。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.8)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.9)$$

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (2.10)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (2.11)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.12)$$

$$h_t = o_t * \tanh(c_t) \quad (2.13)$$

ここで、 W と b はそれぞれ重みとバイアスを表し、 σ はシグモイド活性化関数、 \tanh は tanig 活性化関数を指す。LSTM のこの構造により、長期的な依存関係を効果的にモデル化することが可能となり、特に時系列データや動画予測などの分野において有効である。LSTM の構造の概念図を図 2.4 に示す。

2.2.5 Attention

アテンションメカニズムは、Bahdanau et al. (2014) によってニューラルネットワークをベースにした機械翻訳モデルに導入された。その後、Vaswani et al. (2017) によってアテンションメカニズムを用いた Transformer モデルが提案され、自然言語分野のみならず、画像処理や時系列データなどでも広く用いられるようになった。

アテンションは、モデルが特定の情報に注目することを可能にする機構である。アテンションメカニズムの中心には、クエリ (Query)、キー (Key)、バリュー (Value) の三つの概念がある。これらの要素を使用して、モデルがどの情報に注意を払うべきかを決定する。

- クエリ (Query) : クエリは現在注目している要素や状態を表し、モデルがどの情報に注目するかを決定する基準となる。
- キー (Key) : キーはデータセット内の各要素に関連付けられ、クエリとの関係を定義する。クエリとキーの間の類似性が高いほど、そのキーに関連付けられた情報に注意が向けられる。
- バリュー (Value) : バリューはキーに関連付けられた実際の情報を含み、アテンションメカニズムは、クエリとキーの関係に基づいて、どのバリューを重視するかを決定する。

2.3. 動画予測フレームワーク

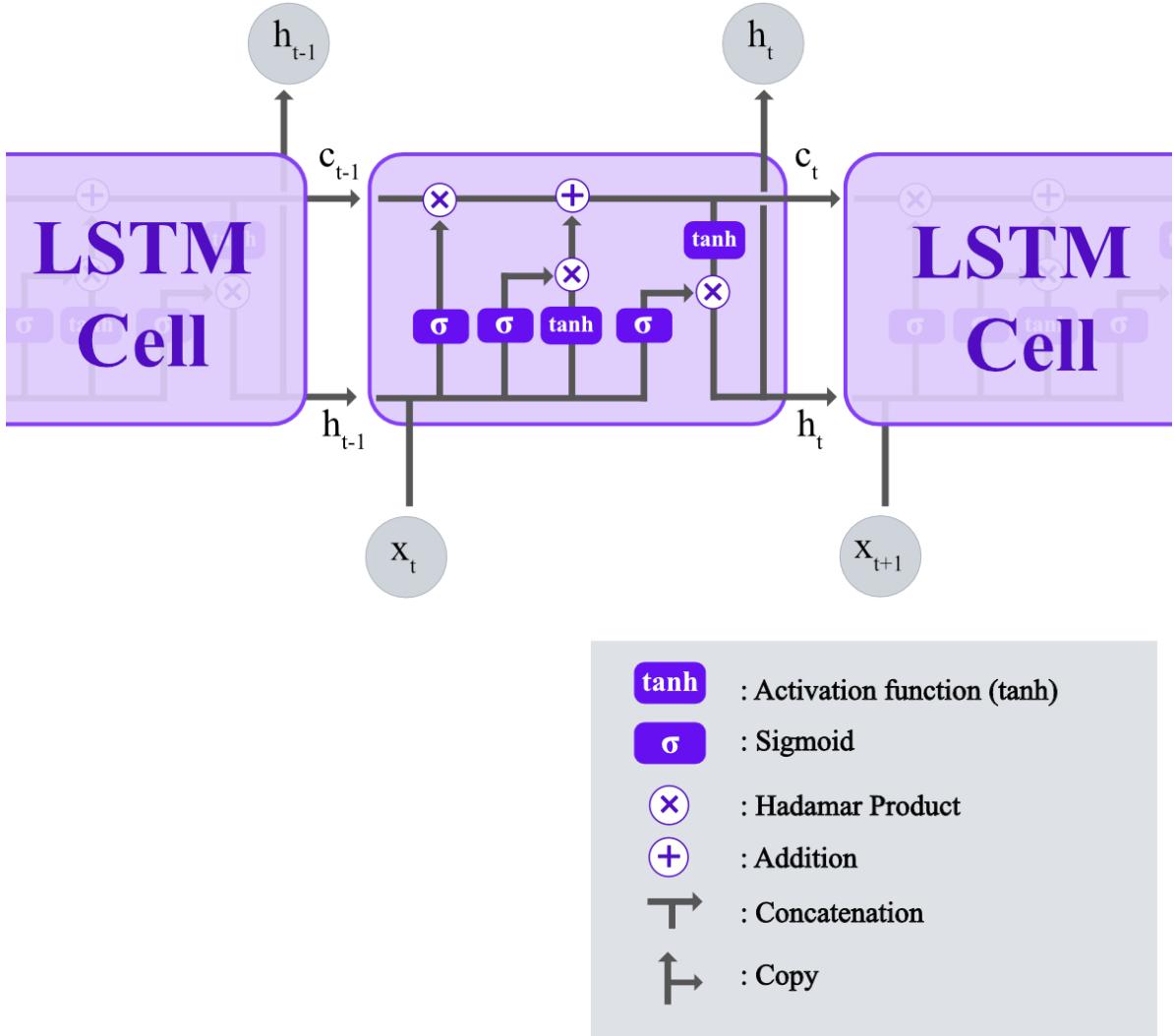


図 2.4: LSTM の構造。RNN のような再帰的な構造を持ちながら、ゲートによって情報の取捨選択を行うことで、長期的な情報の活用が可能になる。

アテンションメカニズムの基本的な操作は、クエリと各キーの間の類似性を計算し、それに基づいて各バリューの重み付き和を取ることである。この重み付き和は次のように定義される。

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2.14)$$

2.3 動画予測フレームワーク

動画予測のフレームワークは、一般的に RNN 構造を基本とし、その中にエンコーダ・デコーダ構造やアテンションメカニズムなどの機構を組み込むことで、動画の時空間的な特徴を効果的に捉

2.3. 動画予測フレームワーク

えることが可能となる。ここでは、もっとも基本的な動画予測フレームワークである ConvLSTM と、その後の改良を加えた PredRNN、また本研究で用いる Motion-Aware Unit (MAU) について説明する。

2.3.1 ConvLSTM

ConvLSTM (Convolutional Long Short-Term Memory) は、Shi et al. (2015) によって提案された、時空間シーケンスデータの処理に特化したニューラルネットワークアーキテクチャである。従来の LSTM の枠組みを拡張し、畳み込み操作を組み込むことで、空間的な情報を効果的に処理する能力を持つ。このような特性により、ConvLSTM は、動画予測のみならず、気象予測や交通流予測など、他の時空間データ処理の応用にも適用可能である。

通常の LSTM では、長期間メモリ c_t やセル出力 h_t に全結合のベクトルを使用するが、ConvLSTM ではそれらに二次元の行列を使用する。また、空間的な特徴をモデル化するために、LSTM の各ゲート（忘却ゲート、入力ゲート、出力ゲート）とセル状態の更新に畳み込み演算を導入する。ConvLSTM のユニットの概念図を図 2.5 に示す。ConvLSTM の数学的定式化は以下の通りである。

$$f_t = \sigma(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \quad (2.15)$$

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \quad (2.16)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_{xc} * X_t + W_{hc} * h_{t-1} + b_c) \quad (2.17)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ c_t + b_o) \quad (2.18)$$

$$h_t = o_t \circ \tanh(c_t) \quad (2.19)$$

ここで、 x_t は時刻 t における入力、 h_t は隠れ状態、 c_t はセル状態を示し、 f_t 、 i_t 、 o_t はそれぞれ忘却ゲート、入力ゲート、出力ゲートの活性化状態を表す。 W と b はネットワークの重みとバイアスパラメータである。 $*$ は畳み込み演算、 \circ はアダマール積（要素ごとの積）を表す。このような計算により、ConvLSTM は時空間データの空間的な特徴と時間的な特徴を統合的に処理し、高度な予測を行う能力を持つ。これにより、モデルは時系列データに含まれる空間的パターンを捉え、それを時間的文脈において解析することが可能となる。

2.3.2 PredRNN

最初の動画予測モデルである ConvLSTM の発表後、そのアーキテクチャを基に様々な改良が加えられてきた。本研究で用いる Motion-Aware Unit の基礎となる PredRNN は、その中でも特に代表的なモデルである。

2.3. 動画予測フレームワーク

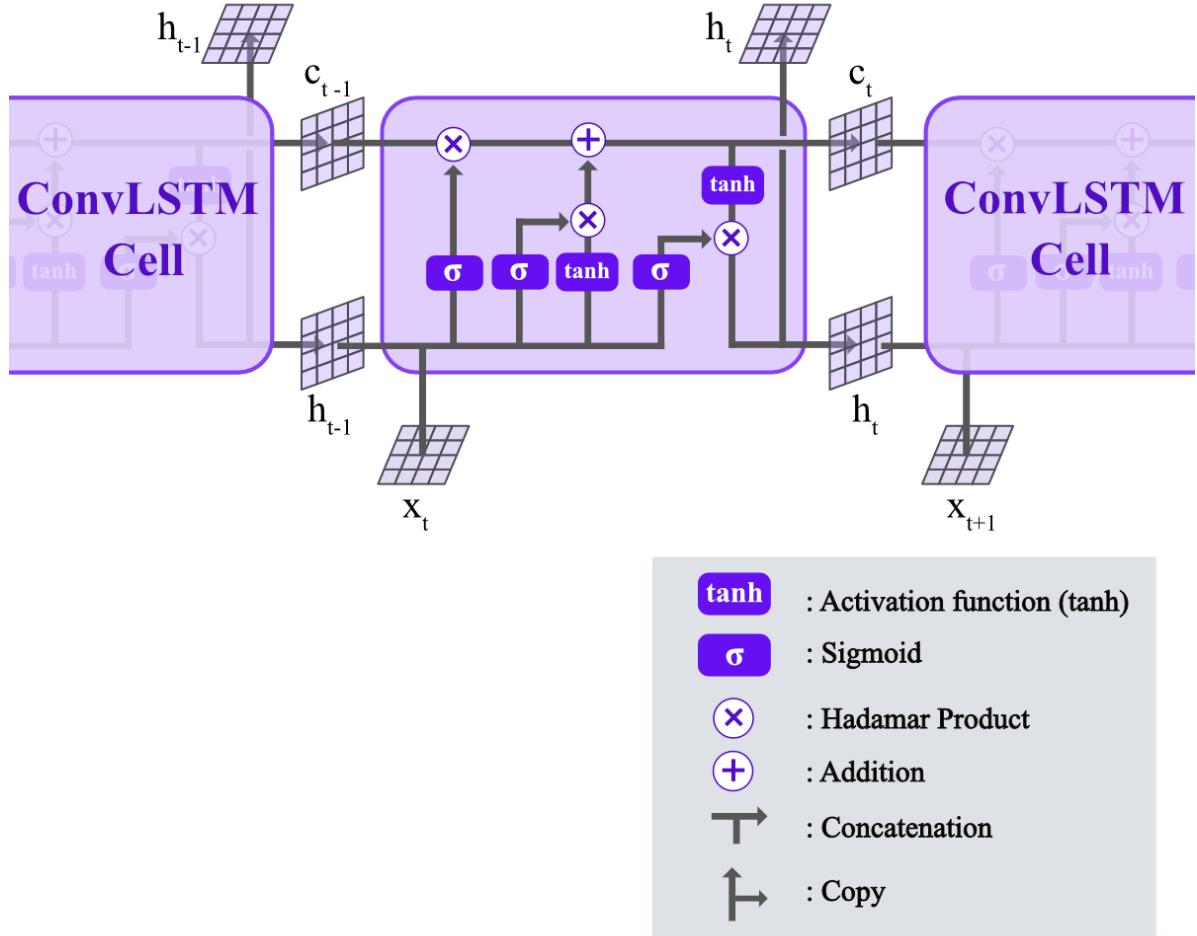


図 2.5: ConvLSTM の構造。

Wang et al. (2017) によって提案された PredRNN は、ConvLSTM を基盤としながらも、いくつかの重要な進化と改良を経て開発された。ConvLSTM はセル状態は各層レベルに対して独立であり、時間方向でのみ更新される。このような状況では、最下層は直前の時間ステップの最上層が生成したセル状態を考慮することができない。PredRNN ではこの概念を拡張し、メモリ状態を異なる層の間で効果的に伝達することを可能にする。

2.3.2.1 時空間メモリフロー

PredRNN では、時空間メモリフローを利用して空間情報の伝達を最適化する。このメモリフローは、遠隔状態間で情報を伝達し、勾配消失問題を軽減するために設計されている。時間方向と、各時間ステップでの隠れ層方向にジグザグにメモリを流す。このように、層をまたいで情報を上方向に伝達し、時間を超えて前方向に情報を伝達することにより、空間情報の効率的な流れを実現し、

2.3. 動画予測フレームワーク

動画フレーム間のより詳細な変化を捉えることができる。

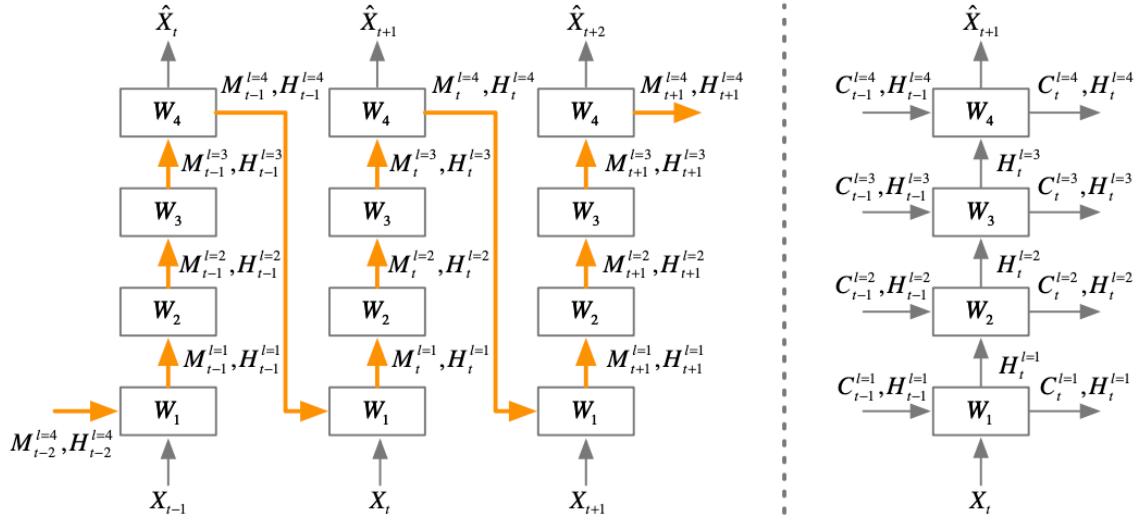


図 2.6: 左が PredRNN における時空間メモリフロー、右が ConvLSTM のメモリフローである。このような異なるレベルの層を通過するメモリフローにより、多様な抽象度の表現を学習することができる。

2.3.2.2 時空間 LSTM ユニット (ST-LSTM)

時空間メモリフローは空間情報の効果的な伝達を可能にするが、水平方向（時間方向）のメモリフローを省略すると、時間的一貫性を犠牲にしてしまう。PredRNN では、標準的な LSTM ユニットを、時空間メモリセルとゲート構造を導入した時空間 LSTM(ST-LSTM) ユニットに置き換える。ST-LSTM ユニットは、時間メモリセルと時空間メモリセルの両方を維持し、それらを縦方向および横方向に流すことにより、一定の時間一貫性を担保しながら、異なる抽象度レベルでの特徴を捉える。また、ST-LSTM は 1×1 の畳み込み層を使用して次元削減を行い、隠れ状態の次元をメモリセルと同じにする。この構造により、PredRNN は時空間データの複雑なダイナミクスを捉え、より正確な予測を生成することができる。

時空間 LSTM の数学的定式化は以下の通りである。ここで、 $*$ は畳み込み演算を、 \circ はアダマール積（要素ごとの積）を示す。 W と b は重みとバイアスパラメータ、 σ はシグモイド活性化関数、 \tanh は双曲線正接活性化関数を指す。 X_t は時刻 t の入力、 H_t^l は隠れ状態、 C_t^l は標準的な

2.3. 動画予測フレームワーク

LSTM セル、 M_t^l は時空間メモリセルを表す。

$$g_t = \tanh(W_{xg} * X_t + W_{hg} * H_{t-1}^l + b_g) \quad (2.20)$$

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1}^l + W_{mi} \circ M_{t-1}^l + b_i) \quad (2.21)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1}^l + W_{mf} \circ M_{t-1}^l + b_f) \quad (2.22)$$

$$C_t^l = f_t \circ C_{t-1}^l + i_t \circ g_t \quad (2.23)$$

$$g'_t = \tanh(W'_{xg} * X_t + W_{mg} * M_{t-1}^l + b'_g) \quad (2.24)$$

$$i'_t = \sigma(W'_{xi} * X_t + W_{mi} * M_{t-1}^l + b'_i) \quad (2.25)$$

$$f'_t = \sigma(W'_{xf} * X_t + W_{mf} * M_{t-1}^l + b'_f) \quad (2.26)$$

$$M_t^l = f'_t \circ M_{t-1}^l + i'_t \circ g'_t \quad (2.27)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1}^l + W_{co} * C_t^l + W_{mo} * M_t^l + b_o) \quad (2.28)$$

$$H_t^l = o_t \circ \tanh(W_{1 \times 1} * [C_t^l, M_t^l]) \quad (2.29)$$

この構造により、PredRNN は時空間データの複雑なダイナミクスを捉え、より正確な未来予測を生成する。

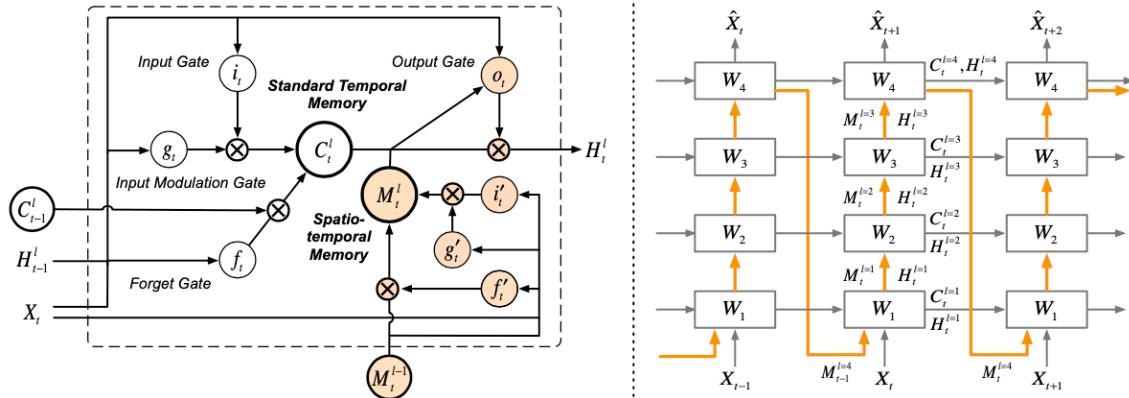


図 2.7: 左が ST-LSTM ユニット、右が PredRNN のメモリフロー。オレンジ色のノードは、従来の ConvLSTM と異なる PredRNN 独自の構造を示す。PredRNN 内のオレンジ色の矢印は、時空間メモリ M_t^l の移行経路を示す。

2.3.3 Motion-Aware Unit(MAU)

Motion-Aware Unit (MAU) は、Chang et al. (2021) によって発表された、フレーム間のダイナミクスをより効率的に捉えるために提案された新しい動画予測アーキテクチャである。MAU は、PredRNN と似た積層 LSTM の構造を持ち、その LSTM ユニットを MAU セルによって置き換

2.3. 動画予測フレームワーク

えている。MAU セルは、注意 (Attention) モジュールと融合 (Fusion) モジュールの 2 つの部分から構成されており、PredRNN における時空間 LSTM ユニットをさらに拡張したものである。このような変更により、時間的受容野

本研究では、この MAU を用いて太陽全球紫外線画像の予測を行う。

2.3.3.1 アーキテクチャ

動画予測モデルでは、出力中の時間ステップが進むほど、時間情報の不確実性が増加するため、予測誤差が劇的に加速してしまう。この問題を解決するため、動画予測モデルはより幅広い時間ステップから有用な特徴を保存し活用する必要がある。すなわち、時間的受容野を拡張する必要がある。このような問題に対するアプローチとして、Wang et al. (2018) によって提案された、三次元畳み込みを導入した E3D-LSTM があったが、非常に高い計算コストを必要とし、性能の改善は限定的であった。Motion-Aware Unit (MAU) は、このような課題を解決するため、Attention 機構を導入した新しいアーキテクチャを提案している。ここでは、Attention 機構を用いたアーキテクチャと、その統合や効率化のために導入されたいいくつかの特徴について説明する。

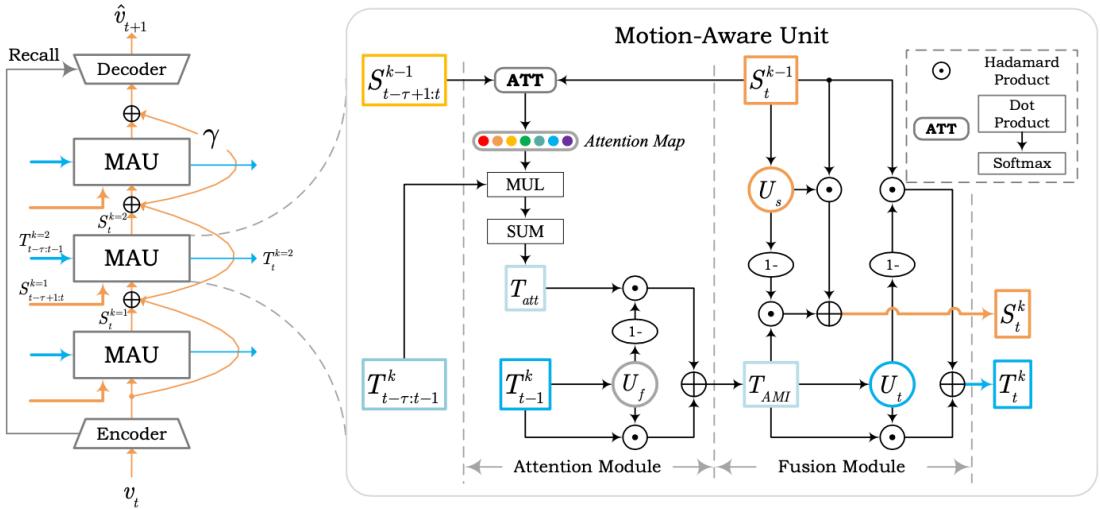


図 2.8: 左: MAU セルを積み重ねた MAU モデルの構造。オレンジ色の矢印は空間方向のメモリフローを示し、青色の矢印は時間方向のメモリフローを示す。エンコーダ・デコーダ構造は各時間ステップにおいて一回ずつ適用され、情報を効果的に圧縮している。右:MAU セルは、注意 (Attention) モジュールと融合 (Fusion) モジュールの 2 つの部分から構成されている。

- エンコーダ・デコーダ構造: MAU は、エンコーダ・デコーダ構造を基本としている。Conv-LSTM や Pred-RNN などの他の動画予測モデルではエンコーダ・デコーダ構造は採用されていないが、MAU ではエンコーダ・デコーダ構造を採用することで、より効率的な特徴抽

2.3. 動画予測フレームワーク

出を可能にしている。一定程度抽象化された情報を再帰的ネットワークに入力することで、より多くの MAU セルの積層を行っても、計算コストを抑えることができる。

$$\hat{X}_t = \text{Dec}[\text{MAU}(\text{Enc}(X_{t-\tau:t-1}))] \quad (2.30)$$

ここで、 Enc はエンコーダ、 Dec はデコーダを表し、 τ は入力シーケンスの長さを表す。

- **注意 (Attention) モジュール:** 注意モジュールは、先述の時間的受容野の拡張のために導入される。このモジュールの導入により、異なる時間状態に対して異なるレベルの注意を払い、予測に対してもっとも相関の高い状態に注意を集中させることができると期待されている。長期的な運動情報としての T_{att} は以下のように計算される。

$$T_{\text{att}} = \sum_{j=1}^{\tau} \alpha_j \cdot T_{t-j}^k \quad (2.31)$$

ここで、 α_j は時間状態 T_{t-j}^k に対するアテンションスコアを表す。 T_{att} は、アテンションスコアによって、予測結果に対して長期的な相関性を持つ時間状態を考慮することができるが、短期的な時間状態 T_{t-1}^k も考慮する必要がある。そこで、それらを融合するためのゲート U_f を導入し、長期的な時間状態 T_{att} と短期的な時間状態 T_{t-1}^k を融合した T_{AMI} を以下のように計算する。

$$U_f = \sigma(W_f * T_{t-k-1}) \quad (2.32)$$

$$T_{\text{AMI}} = U_f \odot T_{t-k-1} + (1 - U_f) \odot T_{\text{att}} \quad (2.33)$$

- **融合 (Fusion) モジュール:** 融合モジュールは、拡張された運動情報 T_{AMI} と現在の空間的状態を適切に統合する役割を持つ。融合プロセスでは、時間的更新ゲート U_t と空間的更新ゲート U_s を導入する。

$$U_t = \sigma(W_{tu} * T_{\text{AMI}}) \quad (2.34)$$

$$U_s = \sigma(W_{su} * X_t) \quad (2.35)$$

ここで、 W_{tu} と W_{su} はそれぞれ時間的更新ゲートと空間的更新ゲートの重みを表す。このゲートを利用して、時間状態 T_t^k と空間状態 S_t^k を以下のように計算する。

$$T_t^k = U_t \odot (W_{tt} * T_{\text{AMI}}) + (1 - U_t) \odot (W_{st} * S_t^{k-1}) \quad (2.36)$$

$$S_t^k = U_s \odot (W_{ss} * S_t^{k-1}) + (1 - U_s) \odot (W_{ts} * T_{\text{AMI}}) + \gamma \cdot S_t^{k-1} \quad (2.37)$$

ここで、 W_{tt} 、 W_{st} 、 W_{ss} 、 W_{ts} はそれぞれ時間状態と空間状態の重みを表し、 γ は学習の安定化を図るための残差項の係数である。

- **情報リコール:** MAU では、エンコーダとデコーダ間で情報損失を防ぐために、情報リコールスキームが採用されている。これは U-Net などで用いられるスキップ接続に似た構造を持っている。これにより、デコーダは多レベルのエンコードされた情報を考慮することができ、予測の視覚的品質を向上させることができる。

2.3. 動画予測フレームワーク

2.3.3.2 主な実験結果

MAU は、複数のデータセットで評価され、その中には Moving MNIST、KITTI、Caltech Pedestrian、TownCentreXVID、Something-Something V2 が含まれる。ここでは、Moving MNIST データセットに関する既存の動画予測モデルによる性能との比較の結果を示す。表 2.1 は、異なる手法による Moving MNIST データセットにおける定量的な結果を示している。MAU は、構造的類似性 (Structual Similarity, SSIM) スコアと平均二乗誤差 (Mean Squared Error, MSE) スコアの両方において、最も優れた結果を示している。2.2 は、LSTM ユニットを積層する動画予測モデルのバックボーンを統一し、それぞれの場合でのパラメータと推論時間に焦点を当てた比較を行なっている。これによれば、MAU は他の手法と比較して、より少ないパラメータ数とより短い推論時間で動作することができ、さらに MSE と SSIM の両方において最も優れた結果を示している。

表 2.1: Moving MNIST データセットにおけるビデオ予測方法の定量的結果 (10 フレーム → 10 フレーム)。低い MSE と高い SSIM スコアはより良い視覚的品質を示す。

方法	SSIM/frame↑	MSE/frame↓
ConvLSTM (NeurIPS2015)	0.707	103.3
FRNN (ECCV2018)	0.819	68.4
VPN (ICML2017)	0.870	70.0
PredRNN (NeurIPS2017)	0.869	56.8
PredRNN++ (ICML2018)	0.898	46.5
MIM (CVPR2019)	0.910	44.2
E3D-LSTM (ICLR2019)	0.910	41.3
CrevNet (ICLR2020)	0.928	38.5
MAU (w/o recalling)	0.931	29.5
MAU	0.937	27.6

2.3. 動画予測フレームワーク

表 2.2: Moving MNIST データセット（10 フレーム→10 フレーム）に対する比較。公平な比較のため、すべてのモデルのエンコーダとデコーダは同じ構造をしており、すべてのモデルは MSE 損失に基づいて Adam オプティマイザーを用いて訓練されている。

方法	バックボーン	MSE↓	SSIM↑	パラメータ数	推論時間
ConvLSTM (NeurIPS2015)	4 × ConvLSTMs	102.1	0.747	0.98M	16.47s
ST-LSTM (NeurIPS2017)	4 × ST-LSTMs	54.5	0.839	1.57M	17.74s
Casual-LSTM (ICML2018)	4 × Casual-LSTMs	46.3	0.899	1.80M	21.25s
MIM (CVPR2019)	4 × MIMs	44.1	0.910	3.03M	45.13s
E3D-LSTM (ICLR2019)	4 × E3D-LSTMs	40.1	0.912	4.70M	57.21s
RPM (ICLR2020)	4 × RPMs	42.0	0.922	1.77M	18.01s
MotionGRU (CVPR2021)	4 × MotionGRUs	34.3	0.928	1.16M	17.58s
MAU	4 × MAUs	29.5	0.931	0.78M	17.34s

第3章

データ

3.1 SDO / AIA

モデルの学習及び評価データとして、NASA の Solar Dynamic Observatory(SDO)(Pesnell et al. [2012]) の Atmospheric Imaging Assembly(AIA)(Lemen et al. [2012]) で撮影された紫外線観測データを用いた。

SDO は NASA の Living With a Star (LWS) プログラムの一つとして 2010 年 2 月に打ち上げられた太陽観測衛星である。AIA、Helioseismic and Magnetic Imager(HMI)、Extreme Ultraviolet Variability Experiment(EVE) などの高い空間解像度、時間分解能を持つ観測機器を搭載し、地上では不可能な多くの波長でのデータを提供する。その観測データを用いることにより、太陽物理学、宇宙天気、また地球環境に関する理解や洞察を深めることができることが期待されている。

AIA は主に太陽大気を観測する観測機器であり、4 つの望遠鏡で構成されている。また、7 つの極紫外線フィルターと、2 つの紫外線フィルター、および 1 つの可視光フィルターを持ち、広範な温度帯で太陽大気を観察することを可能にしている。本研究で用いられる 171 Å、193 Å、211 Å の 3 つのフィルターは、36 秒間隔で撮影され、 4096×4096 、約 1.5 秒角の空間解像度を持つ

これらのデータは Joint Science Operations Center(JSOC) によって提供されており、Python の太陽物理学を支援するライブラリである Sunpy を用いてダウンロードすることができる。

3.1.1 AIA 211 Å

表 3.1 のように、211 Å (21.1 nm) のフィルターは、約 200 万 K の 14 倍鉄 (Fe XIV) イオンが放射するスペクトルを捉えるために特化している。この波長での観測は、活動領域のコロナを観測するのに最適である。図 3.1 は、これらの特性を示している。

3.1. SDO / AIA

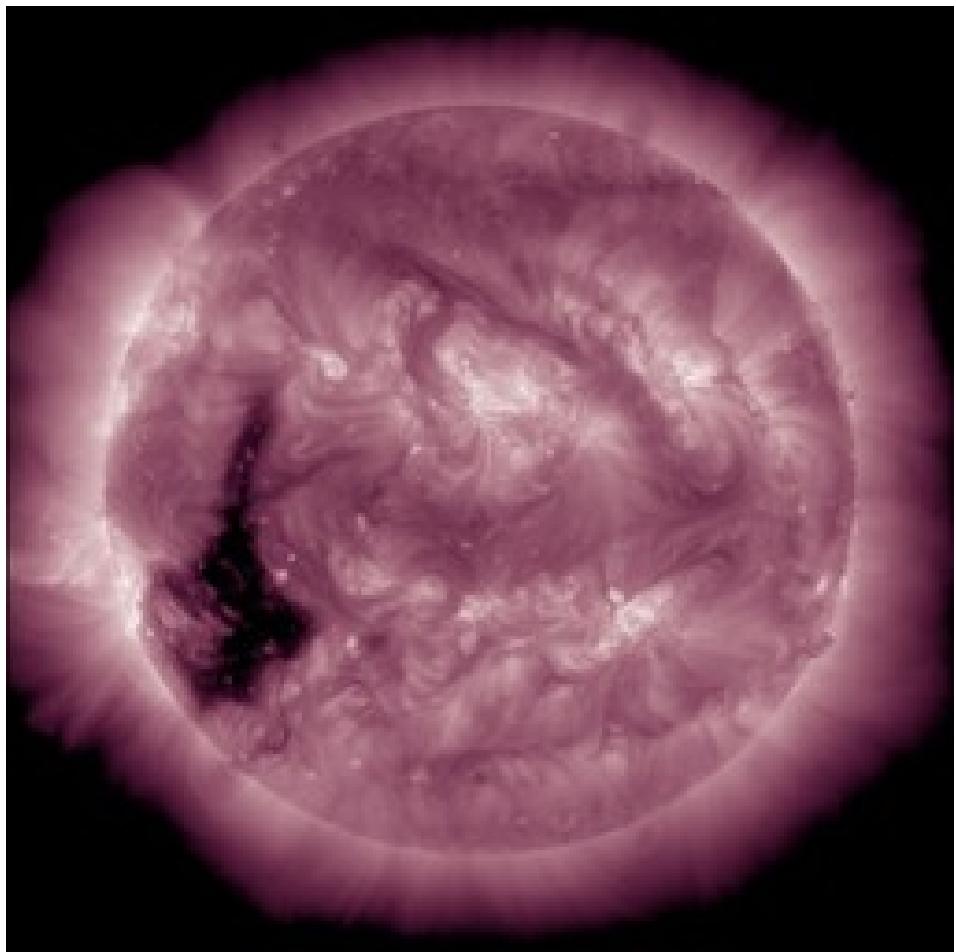


図 3.1: SDO/AIA の 211 Å フィルターで撮影された太陽全球紫外線像。強調のために紫色に色付けされている。球面の中上部から中下部には明るく輝く活動領域が見られ、左下部に暗くコロナホールが観測できる。

3.1.2 AIA 193 Å

表 3.1 のように、193 Å (19.3 nm) のフィルターは、約 150 万 K の 12 倍鉄 (Fe XII) イオンが放射するスペクトル、または約 2000 万 K の 24 倍鉄 (Fe XXIV) イオンが放射するスペクトルを捉えるために特化している。前者は主にコロナの中程度の高温領域を観測するために用いられ、後者は、主にコロナの高温フレアプラズマを観測するために用いられる。さらに、コロナホールも強調して観測することができる。図 3.2 は 193 Å フィルターで捉えた太陽像を示している。

3.1. SDO / AIA

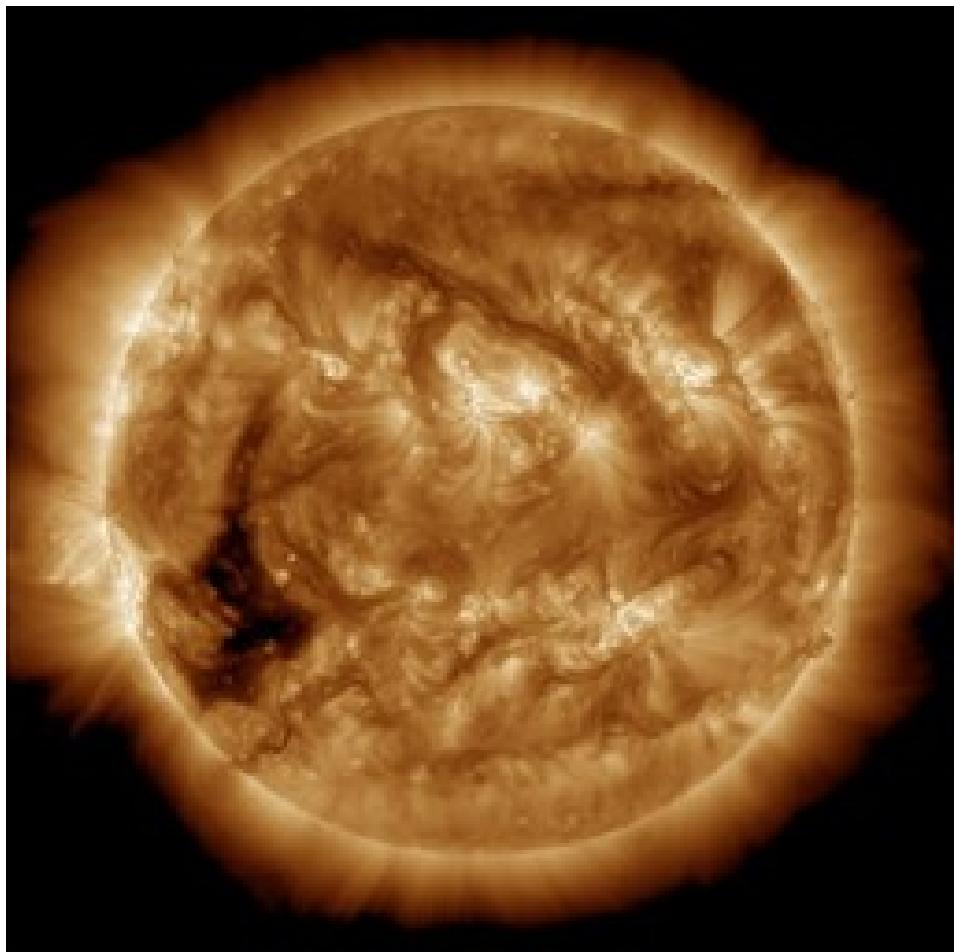


図 3.2: SDO/AIA の 193 Å フィルターで撮影された太陽全球紫外線像。強調のために橙色に色付けされている。

3.1.3 AIA 171 Å

表 3.1 のように、171 Å (17.1 nm) のフィルターは、約 60 万 K の 9 倍鉄 (Fe IX) イオンが放射するスペクトルを捉えるために特化している。この波長での観測は、太陽のコロナループ、静穏領域コロナ、コロナホールなどの磁気構造を詳細に観察することができる。図 3.3 に示される 171 Å フィルターによる観測は、これらの特徴を捉えている。

3.2. 前処理

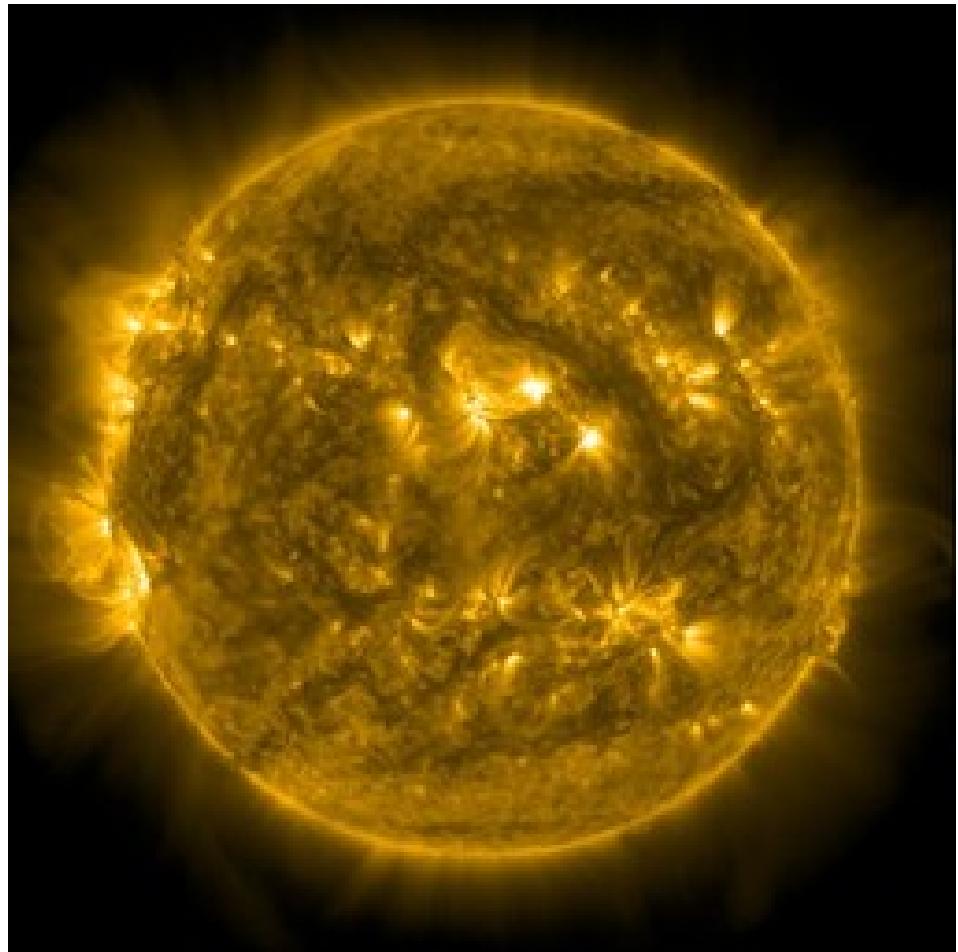


図 3.3: SDO/AIA の 171 Å フィルターで撮影された太陽全球紫外線像。強調のために黄色に色付けされている。193 Å、211 Å では観察できない、コロナホールなどの静穏領域のスペクトルも観測できる。

表 3.1: SDO AIA の 171, 193, 211(Å) の特性

フィルター (Å)	主要イオン	観測される主要な太陽活動または領域	温度帯 (K)
171	Fe IX	静穏領域コロナ、上層遷移領域	6.3×10^5
193	Fe XII, XXIV	コロナと高温フレアプラズマ	$1.5 \times 10^6, 2.0 \times 10^7$
211	Fe XIV	活動領域コロナ	2.0×10^6

3.2 前処理

本研究で用いるデータセットには、SDO/AIA のデータが提供されている 2010 年 5 月から、2022 年 10 月までのデータが含まれている。この期間に存在するデータから、4 時間ごとにデータを抽

3.2. 前処理

出し、各波長ごとに約 22000 枚をデータセットに含んでいる。データは JSOC により提供されているものをダウンロードし、その後、不正な画像を除去し、正規化やスケーリングを行った後、学習用、検証用、テスト用に分割した。この手順をアルゴリズム 1 に示す。これらのデータを、24 枚の画像を 1 セットとして分割する。各セットは 24 枚の時系列に並んだ画像で構成され、太陽全球の空間的情報の時間的変化を捉えている。24 枚のうち、前半の 12 枚、すなわち 48 時間までを入力シークエンス、後半の 12 枚、すなわち 52 時間から 96 時間までを出力シークエンスとして扱う。学習の際は、入力シークエンスに対して出力シークエンスを教師データとして扱い、テストの際は入力シークエンスに続くモデルにとって未知の出力シークエンスを再現できるか検証する。

このデータセットは第 24 太陽活動周期の初期から、第 25 周期の初期までの観測データを網羅している。この時間範囲には、太陽活動の活発性が高いフェーズと低いフェーズの両方が含まれている。従って、このデータセットは太陽活動の活発性に依存しない可能性が高く、その汎化能力に対する期待が一定程度裏付けられる。

3.2.1 不正な画像の除去

SDO/AIA 望遠鏡で撮影された全球画像には、露光時間が他の画像より極端に低い、画像内に太陽全体を捉えていない、などの不正な画像が含まれている。確認することができた主な不正な画像を図 3.4 に示す。

これらの画像は、モデルの学習に悪影響を及ぼす可能性があるため、データセットから除去した。機械学習のタスクによっては、十分なデータセットがあれば、モデルが不正な画像に対する頑健性を獲得し、不正な画像がデータセットに含まれていても、学習結果にあまり大きな影響を与えない場合がある。しかし、本研究で行う動画予測は、データセットに含まれる画像がそのまま教師データとなるため、不正な画像は損失関数の計算、またはモデルの評価に大きな影響を与えるため、慎重に除去する必要がある。データの除去には、FITS ファイルのヘッダーに記録された各キーワードの値に対して閾値を設定して判定したのち、numpy による輪郭検出を用いた月蝕判定関数により不正な画像を排除した。

3.2.2 スケーリングと正規化

太陽動画データセットの前処理として、以下のステップを実施する。

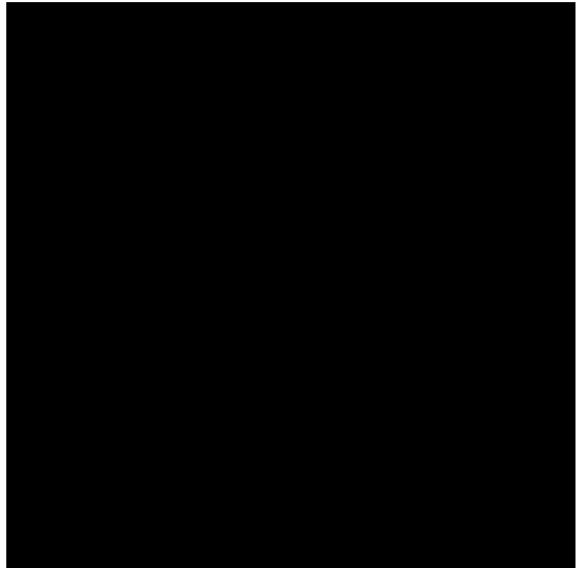
1. 正規化: クリッピング処理されたデータを 0 から 1 の範囲に正規化する。ここで、ノイズによる負の値を削除し、極端に大きい外れ値の影響を削減するために、画像内の全ピクセル値に対して

3.2. 前処理

Algorithm 1 データセット作成アルゴリズム

```
1: procedure CREATESOLARDATASET
2:   for each wavelength in wavelengths do
3:     images  $\leftarrow$  DOWNLOADDATA(wavelength)
4:     images  $\leftarrow$  VALIDATEANDREPLACEIMAGES(images)
5:     all_images.extend(images)
6:   end for
7:   dataset  $\leftarrow$  CREATEDATASET(all_images)
8:   processed_dataset  $\leftarrow$  PREPROCESSDATASET(dataset)
9:   train, val, test  $\leftarrow$  SPLITDATASET(processed_dataset)
10:  return train, val, test
11: end procedure
12: function DOWNLOADDATA(wavelength)
13:   Download data for given wavelength
14:   return images
15: end function
16: function VALIDATEANDREPLACEIMAGES(images)
17:   for each image in images do
18:     if not VALIDATEIMAGE(image) then
19:       alternative  $\leftarrow$  FINDALTERNATIVEIMAGE(image.timestamp)
20:       images.replace(image, alternative)
21:       VALIDATEANDREPLACEIMAGES(images)                                 $\triangleright$  Recursive call
22:     end if
23:   end for
24:   return images
25: end function
26: function CREATEDATASET(images)
27:   Create dataset by grouping 24 images
28:   return dataset
29: end function
30: function PREPROCESSDATASET(dataset)
31:   Apply preprocessing to each image in dataset
32:   return processed_dataset
33: end function
34: function SPLITDATASET(dataset)
35:   Split the dataset into train, validation, and test sets
36:   return train, val, test
37: end function
```

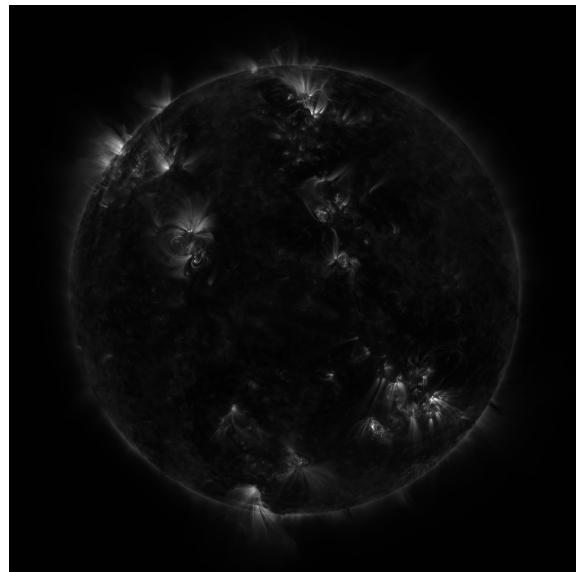
3.2. 前処理



(a) 短い露光時間により、極端に暗い画像。



(b) 太陽が画像の中心にない画像。



(c) 衛星が回転しており、正しい角度で太陽が撮影されていない画像。活動領域の少ない左下部と右上部が極である。

図 3.4: SDO/AIA により観測された不正な画像の例

3.2. 前処理

最小値を 0、最大値を 10000 に設定した:

$$I_{normalized}(x, y) = \frac{\min(\max(I(x, y), 0), 10000)}{10000} \quad (3.1)$$

2. 平方根スケーリング: ダイナミックレンジの広さに対応するために、正規化されたデータに平方根スケーリングを適用する。この過程は以下の式で示される。

$$I_{scaled}(x, y) = \sqrt{I_{normalized}(x, y)} \quad (3.2)$$

図 3.5 のヒストグラムに示すように、スケーリングが適用されていないデータ（左）は、下位 5% 程度の範囲に極端に輝度が集中している。平方根スケーリングを適用したデータ（右）では、そのような極端な輝度の差が緩和されている。

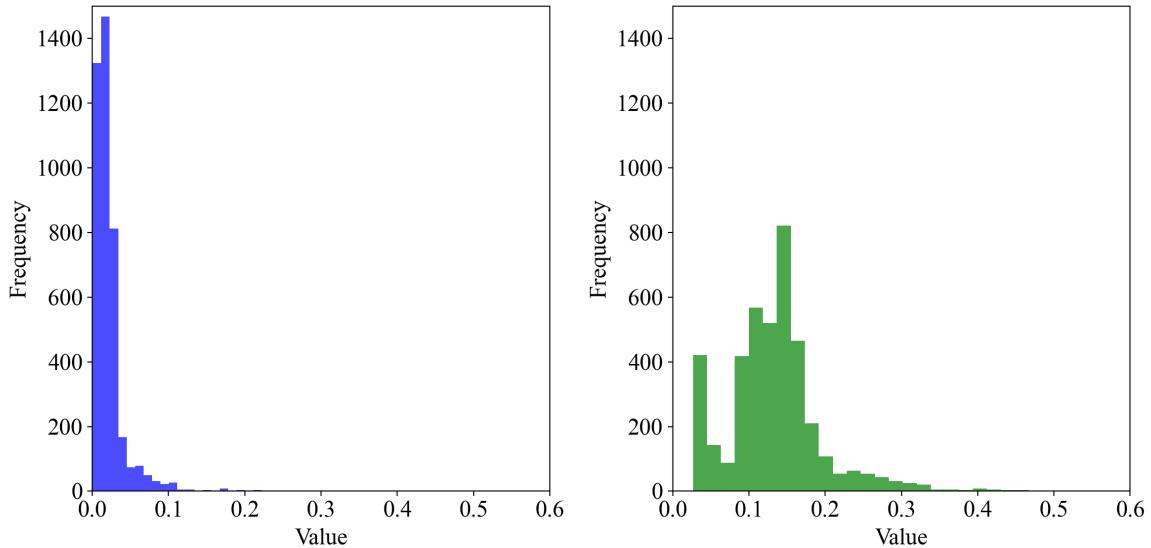


図 3.5: ある画像での正規化されたデータに対する平方根スケーリングの効果。左: 正規化されたデータのヒストグラム。右: 平方根スケーリングを適用したデータのヒストグラム。

3. リサイズ: 効率的な処理のために、4192px × 4192px の画像を 512px × 512px の解像度にリサイズした。この処理は、画像の空間解像度を低下させるが、太陽全球の大規模な構造を捉えるには十分である。

3.2.3 データセットの分割

このようにして作成されたデータセットは、約 1000 セットになり、これを学習用データセットに約 800 セット、検証データセットに 50 セット、テストデータセットに 50 セットというように分割した。データセット 1 単位は、24 枚の画像から構成される。

3.2. 前処理

実験 1 では、211 Å フィルターの画像のみを用いた。実験 2 では、171 Å、193 Å、211 Å の 3 つのフィルターの画像を用いた。最終的な分割とデータセットの概要を表 3.2 に示す。

実験	実験 1	実験 2
入力波長	211 Å	171 Å, 193 Å, 211 Å
出力波長		211 Å
総枚数	22000	66000
セット数		約 1000
セットごとの枚数		入力 12 → 出力 12
解像度		512 * 512

表 3.2: 各実験でのデータセット

第 4 章

Motion-Aware Unit を用いた 1 波長を入力とした紫外線像の全球時系列予測

4.1 実験概要

この実験では入力、出力ともに 211 Å フィルターで得られたデータを利用した。これは 211 Å フィルターで撮影された紫外線像が、コロナホールと活動領域といった、二つの太陽円盤上の大規模構造をバランスよく明瞭に表現し、本研究のモデルの効果検証に適していると考えたためである。モデルには Motion-Aware Unit を用い、1 波長のデータを入力として、全球の時系列予測を行った。この実験の概要を図 4.1 に示す。

4.2 実験設定

各ハイパーパラメータの設定を表 4.1 に示す。バッチサイズは実験的に決定し、最も安定的に最終的に良好な精度を達成できた値を採用した。また、エポック数は 100 とした。学習率は 0.0005 とした。MAU Cell 数は、Chang et al. [2021] の実験設定を参考に、16 とした。学習時間の短縮およびメモリ使用量の削減のため、学習時には Automatic Mixed Precision (AMP)(Micikevicius et al. [2017]) を用いた。これは、単精度浮動小数点演算と半精度浮動小数点演算を適切に混在させることで、モデル性能をほとんど落とさずに計算資源を節約し学習を高速化する手法である。GPU は NVIDIA RTX A6000 を用いた。

4.3. 学習の推移

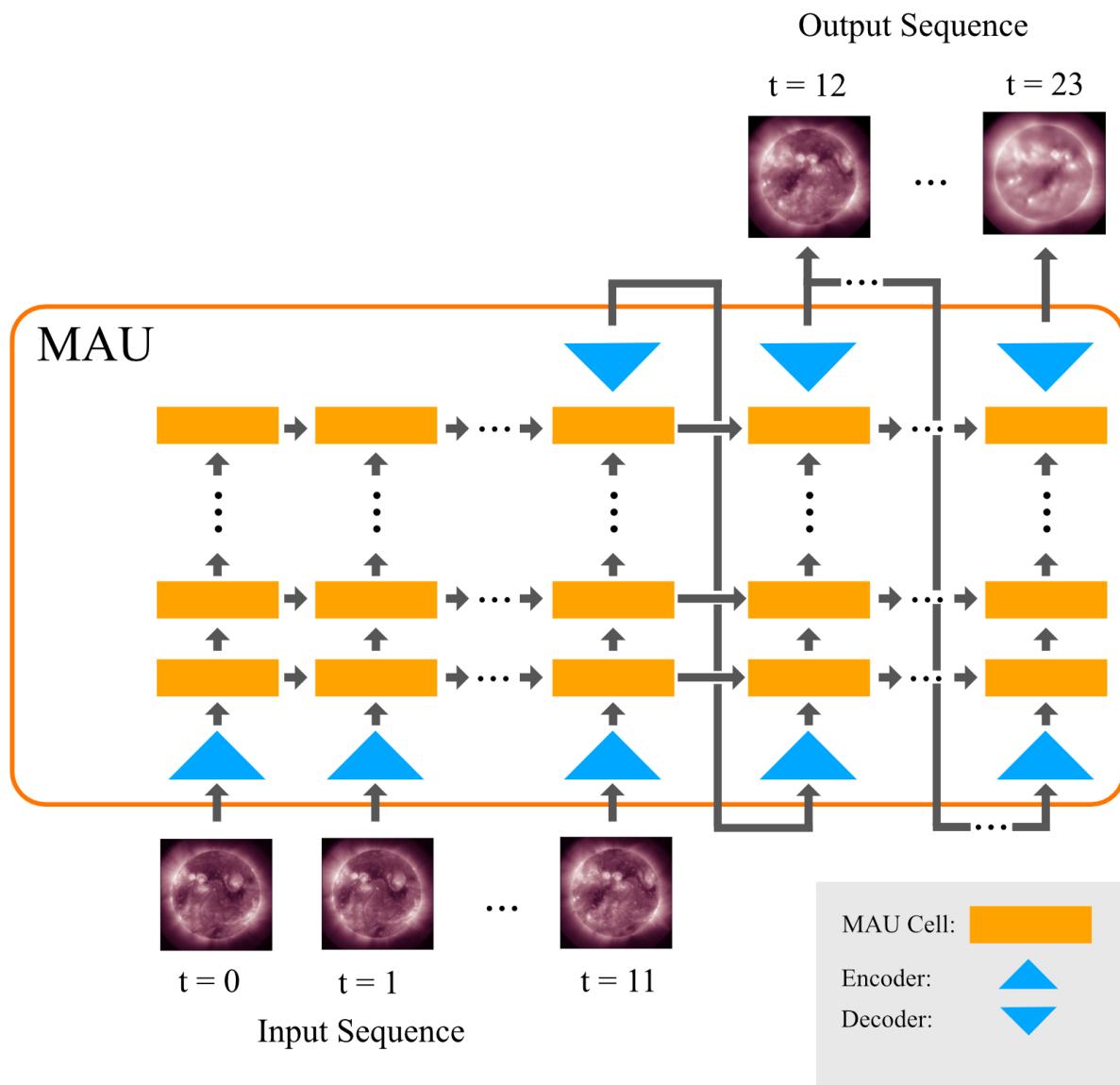


図 4.1: 実験の概念図。モデルには Motion-Aware Unit を用い、1 波長のデータを入力として、全球の時系列予測を行った。

4.3 学習の推移

学習は図 4.2 のように推移した。学習損失は全体的に安定して推移し、検証損失は時折急激に値が増加しているが、全体的には減少している。学習の完了までには約 12 時間を要した。

4.4. 実験結果

ハイパーパラメータ	値
バッチサイズ	4
エポック数	100
学習率	0.0005
損失関数	MSE
チャンネル	1
カーネルサイズ	(5, 5)
MAU Cell 数	16

表 4.1: 本実験でのハイパーパラメータ設定

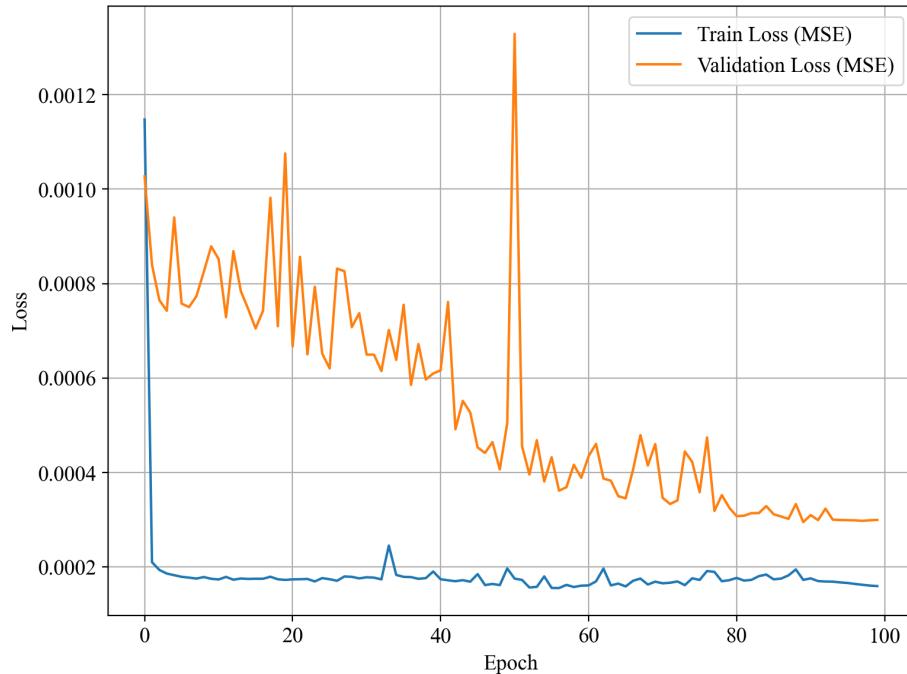


図 4.2: 本実験での、学習データ、検証データでの損失関数の推移。学習の損失は安定している。検証の損失は振動しながら減少している。

4.4 実験結果

図 4.3 および図 4.4 に、この実験での出力例を示す。これは学習データに含まれない期間のテストデータである。 モデルの出力は、視覚的には実際の観測画像と概ね合致しており、特に自転による大規模構造の移動といった顕著な時間的特徴は再現できていることがわかる。

4.4. 実験結果

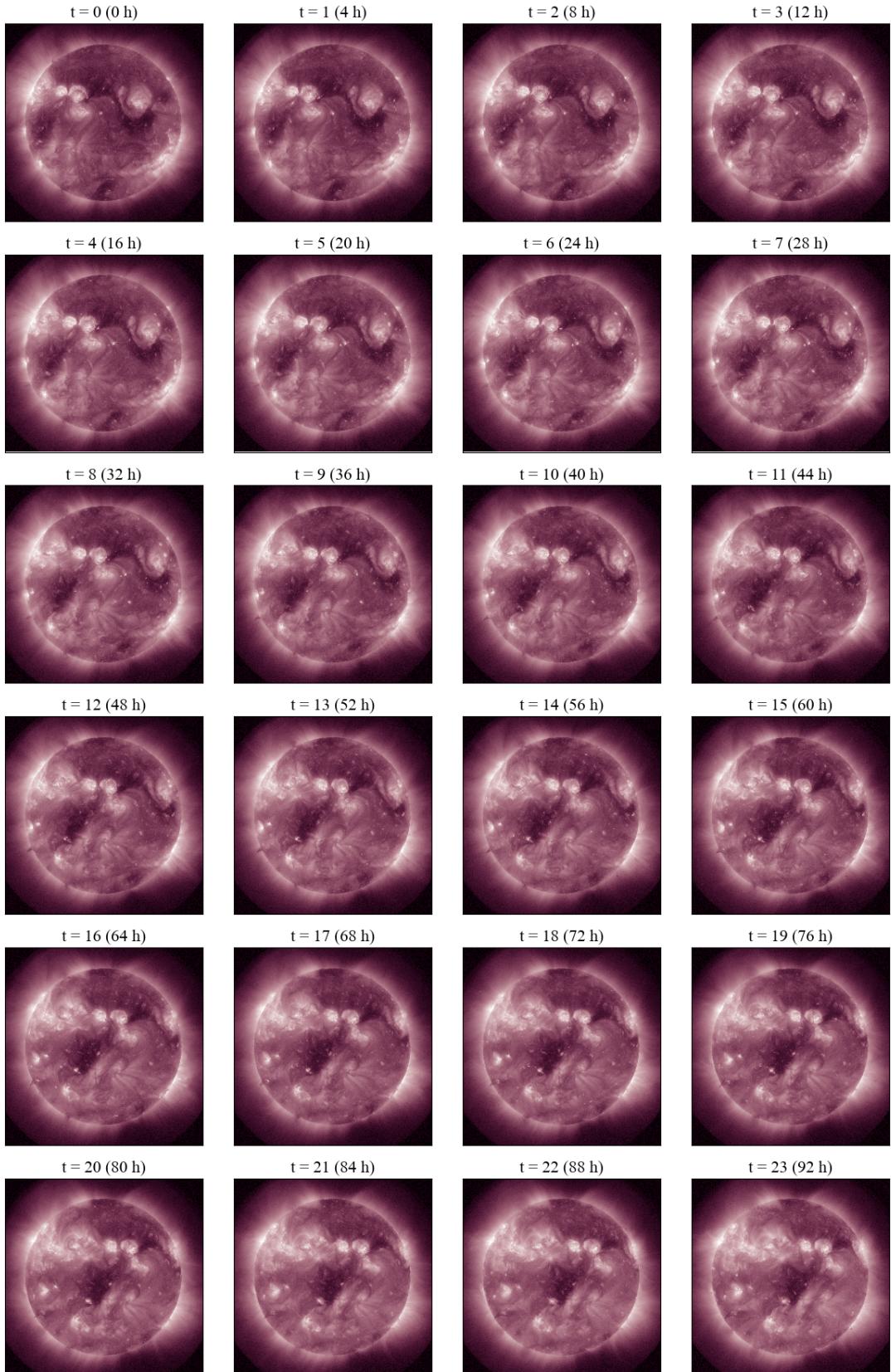


図 4.3: 実際の観測画像の例。2022 年 10 月 28 日 0 時から 2022 年 11 月 1 日 20 時までの期間から 4 時間毎にサンプリングされている。この $t=3$ から $t=11$ までをモデルに入力データとして渡している。モデルはその入力データを元に、 $t=12$ から $t=23$ の 12 枚の画像を予測する。 $t=12$ 以降の実際の観測画像はモデルに渡されない。

4.4. 実験結果

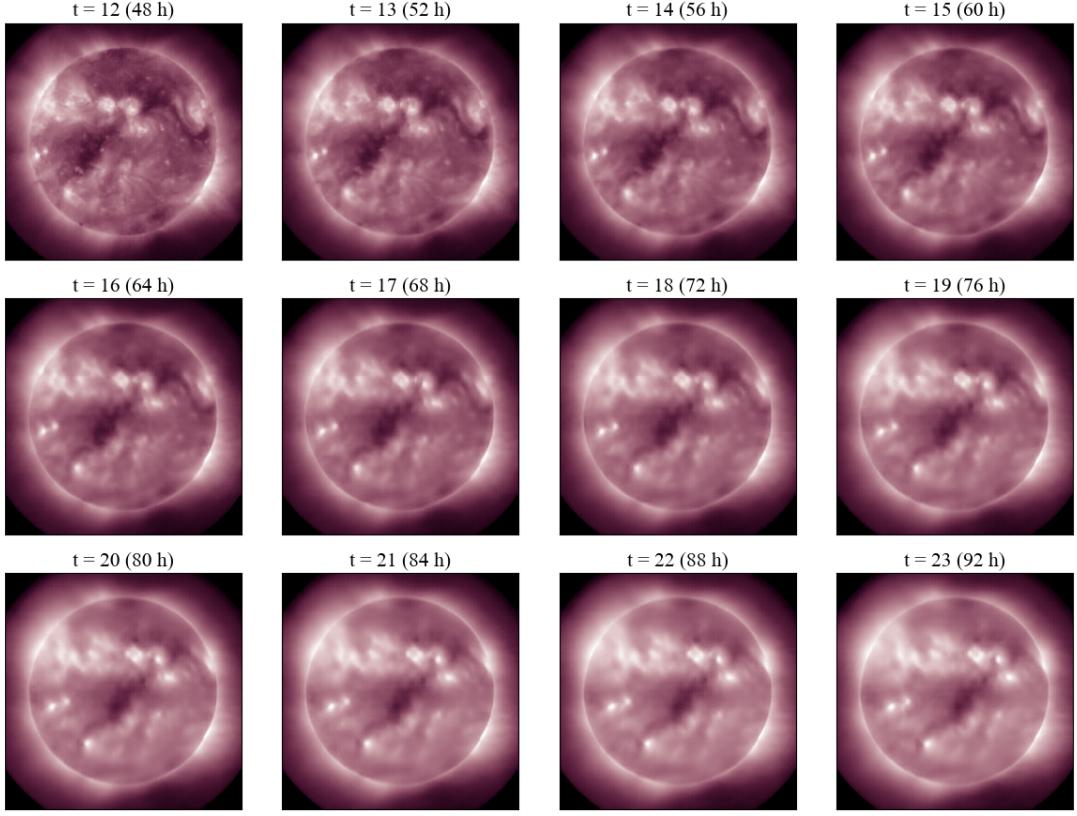


図 4.4: MAU による予測画像。対応するタイムステップ t の観測画像 (図 4.3) と比較することでモデルの再現度を視覚的に評価することができる。大規模な構造は概ね実際の観測画像と合致している。モデルの特性により、時間経過とともに少しずつ予測が不安定になり、ぼやけた見た目になる。

動画予測の精度を評価するために、太陽の輝度強度の再現性を定量的、またまたは視覚的に評価する。これは、Nishizuka et al. (2018) や～～などの、太陽画像から太陽イベントを予測する先行研究では、その画像中の輝度強度を主要な特徴量として採用していることに基づく。この輝度強度の再現性の評価を、さまざまな条件下で行った。はじめに全球での評価を行い、次に経度依存性の評価を行った。最後に、東側リムから出現する活動領域に対する視覚的評価を行った。

4.4.1 全球での評価

はじめに全球での評価を行った。この評価では、まず輝度強度の平均値と実際の平均値との誤差、構造的類似度 (Structual Similarity, SSIM) を計算した。さらに単純差動回転モデルとの比較も行った。これらの値の時間経過に対する変化を観察し、より不確定性の高い将来の予測に対しても動画予測モデルが有効であるかを検証した。

4.4. 実験結果

4.4.1.1 平均輝度の再現

■平均輝度の絶対誤差の計算 テストセット全体における、ある時間ステップ t の平均輝度の絶対誤差を以下のように計算した。

$$\bar{E}_t = \frac{1}{N} \sum_{i=1}^{50} |\bar{I}_{\text{Prediction}_{i,t}} - \bar{I}_{\text{Actual}_{i,t}}| \quad (4.1)$$

ここで、 i はテストセットのインデックスを表す。また、 $\bar{I}_{\text{Prediction}_{i,t}}$ は、テストセット i 、時間ステップ t における、モデルから生成された画像から計算された平均輝度を表し、 $\bar{I}_{\text{Actual}_{i,t}}$ は、実際の画像から計算された平均輝度を表す。平均輝度は全球 (画像中の太陽の球面) に対してのみ行い、画像中の背景や外縁部からはみ出すコロナなどはその計算に含まれない。背景から全球に対して切り出される部分は、図 4.5 に示されている。この全球の定義および計算は、取得した FITS ファイルのヘッダーに記載される太陽の中心および半径に基づいている。

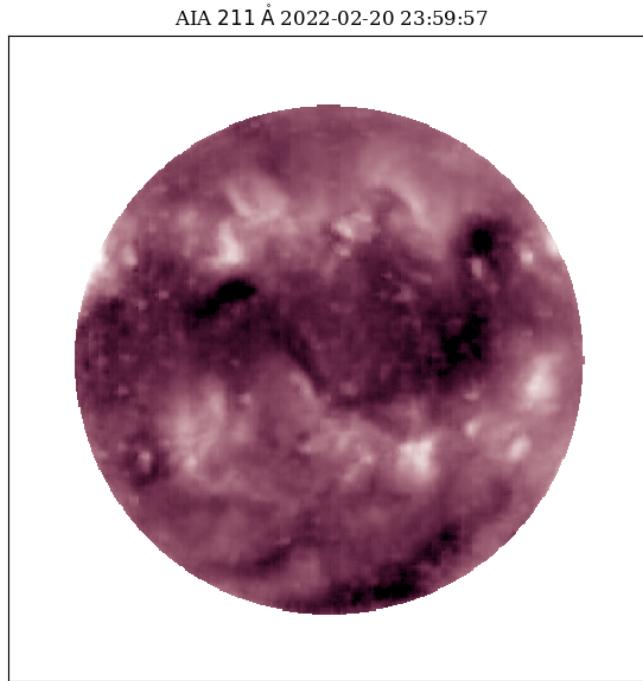


図 4.5: 生成した画像から全球部分のみ切り出した画像の例。この部分にのみ平均輝度を計算する。

モデルの出力の全球での平均輝度と、実際の観測画像との誤差の推移を図 4.6 に示す。これは、50 のテストセットに対して、各テストセットに含まれる各画像の全球での平均輝度を計算し、その時間ステップごとの平均値を取ったものである。輝度の推移のみから特定の傾向を見出すことは難しいが、全体として平均絶対誤差は 4% 以下に収まっている。

4.4. 実験結果

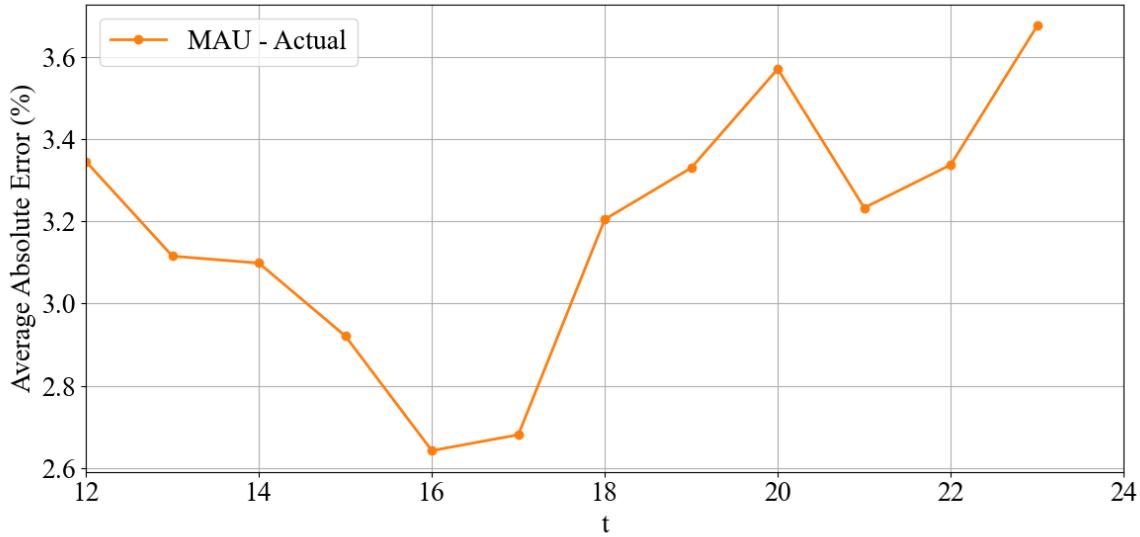


図 4.6: MAU によるテストセットの予測画像と実際の観測画像の平均絶対誤差の時間推移。横軸が時間ステップ、縦軸が平均絶対誤差を表す。時間ステップを経る毎に誤差は単調に上昇していくが、その誤差は最終タイムステップでも 5% 以下にとどまっている。

さらに、入力シークエンスの最後から 48 時間後の画像の全球での平均輝度と、実際の観測画像との差異を観察する。その散布図 4.7 に示す。このタイムステップは、出力の最後のタイムステップであり、最も不確定性の高い予測である。相関係数は 0.97 であり、非常に良好な値である。実際の観測画像の平均輝度の高低に関わらず、高い精度で平均輝度を再現できていることがわかる。

■単純差動回転モデルとの比較 モデルの予測性能をさらに詳細に評価するために、シンプルな差動回転モデルとの比較を行った。目視や、平均輝度から、モデルの出力は実際の観測画像と概ね合致しており、特に自転による構造的変化などの主要な時間的特徴を再現できていることがわかった。ここでは、単純差動回転シミュレーションモデルによる出力と、我々のモデルの出力の再現精度の比較を行う。これにより、モデルが単に自転を予測しているのではなく、より複雑な時間的変化を予測できているかを検証した。

太陽は自転するが、その実体は流体であるため、緯度によって自転速度が異なる。極付近の自転周期は約 35 日であるが、赤道付近では約 25 日である。この現象を差動回転と呼ぶ。差動回転をシミュレーションする研究は盛んに行われているが、ここでは回転速度を緯度依存としてモデル化する Howard et al. (1990) Howard et al. [1990] の差動回転モデルを用いた。このモデルは、Heliographic 緯度 θ に対する回転速度 $\omega(\theta)$ を以下のように定義する:

4.4. 実験結果

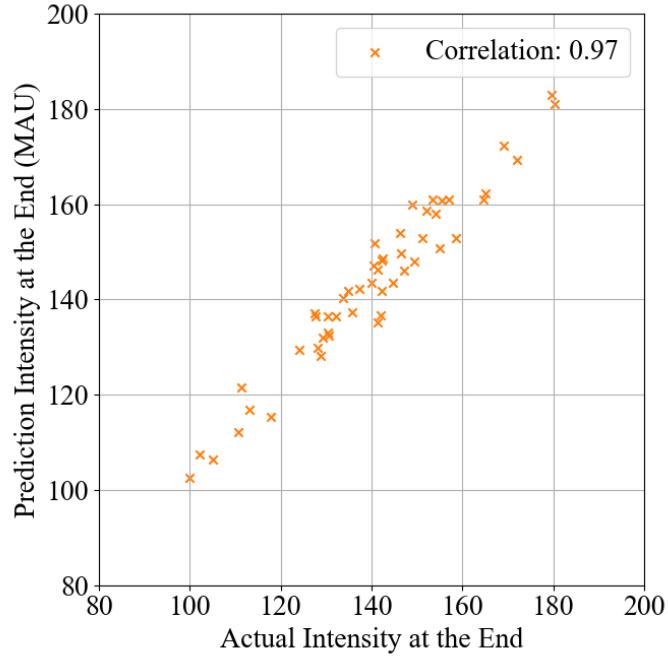


図 4.7: テストセットの最終ステップにおける全球平均輝度の予測対実測の散布図。縦軸が MAU による予測から計算された平均輝度強度、横軸が実際の観測画像から計算された平均輝度強度を表す。計算された相関係数は 0.97 である。

$$\omega(\theta) = A + B \sin^2(\theta) + C \sin^4(\theta) \quad (4.2)$$

$$\text{where } A = 2.894 \mu\text{rad/s}, \quad (4.3)$$

$$B = -0.428 \mu\text{rad/s}, \quad (4.4)$$

$$C = -0.370 \mu\text{rad/s} \quad (4.5)$$

このモデルは Sunpy に実装されており、*physics.differential_rotation* というモジュールとして提供されている。このモジュールによる画像の生成は、全球の各ピクセルに対して、そのピクセルの緯度に対する回転速度を計算し、その速度で西に向かって各ピクセルを移動させることで行われる。この単純差動回転モデルによるシミュレーションの例を図 4.8 に示す。このピクセルの移動による画像の生成は、全球面以外の背景には計算されない。また、東側外縁部からは入力がない新しい太陽表面が出現するためん、この部分は計算されない。

比較は、単純差動回転モデルによるシミュレーションと、実際の観測画像との平均輝度の絶対誤差を計算し、それを前述の動画予測によるものと比較することで行った。この誤差の推移を図 4.9 に示す。単純差動回転モデルは時間経過とともに誤差が単調に増加するが、動画予測モデルは最終タイムステップにおいても誤差が 4% 以下に収まっている。t=12 から 16 の間は、単純差動回転モ

4.4. 実験結果

ルの方が誤差が小さいが、それ以降は逆転し、動画予測モデルの方が誤差が小さくなっている。このことから、動画予測モデルは、単純差動回転モデルよりも、より時間経過に対して堅牢であることがわかる。

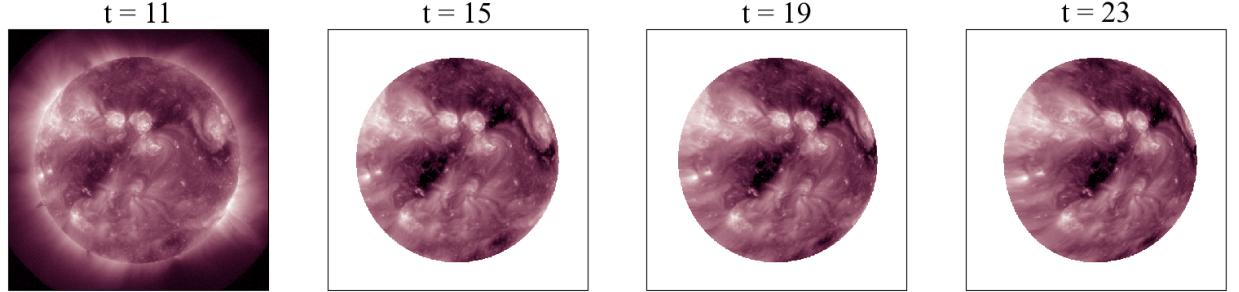


図 4.8: Howard (1990) による差動回転モデルによるシミュレーションの例。入力シークエンスの最終入力 ($t=11$)(左) をもとに、各ピクセルに式 (4.2) を適用し移動させることで画像を生成する。全球面以外の背景には計算されないため、徐々に東側から画像が欠けていくのがわかる。

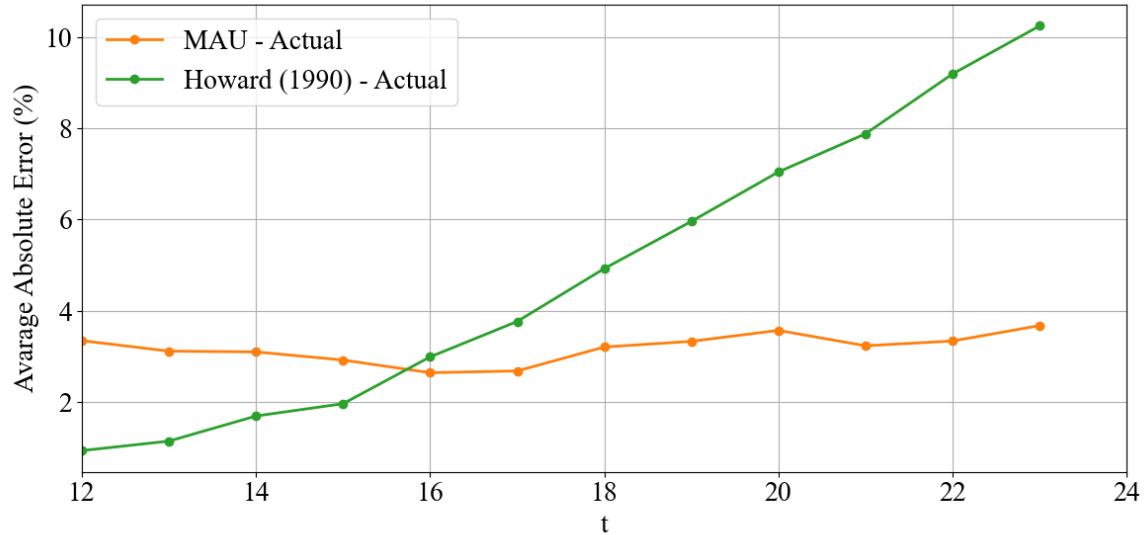


図 4.9: MAU によるテストセットの予測画像と実際の観測画像の平均絶対誤差 (オレンジ) と、単純差動回転モデルと実際の観測画像の平均絶対誤差 (緑)。

さらに、出力シークエンスの最後のタイムステップにおいて、単純差動回転モデルによるシミュレーションと、実際の観測画像との差異を観察し、動画予測モデルによる出力と比較した。その散布図 4.10 に示す。単純差動回転モデルによるシミュレーションの平均輝度と、実際の観測画像の平均輝度は、相関係数では 0.85 である。データ点が全体的に左上に偏っていることから、単純差動回転モデルは、実際の観測画像よりも平均輝度を高く予測していることがわかる。一方で、前述

4.4. 実験結果

のように、動画予測モデルによる出力の平均輝度と、実際の観測画像の平均輝度は、相関係数では 0.97 である。ここから、やはり最終タイムステップでの予測においても、動画予測モデルは単純差動回転モデルよりも高い精度で平均輝度を再現できていることがわかる。

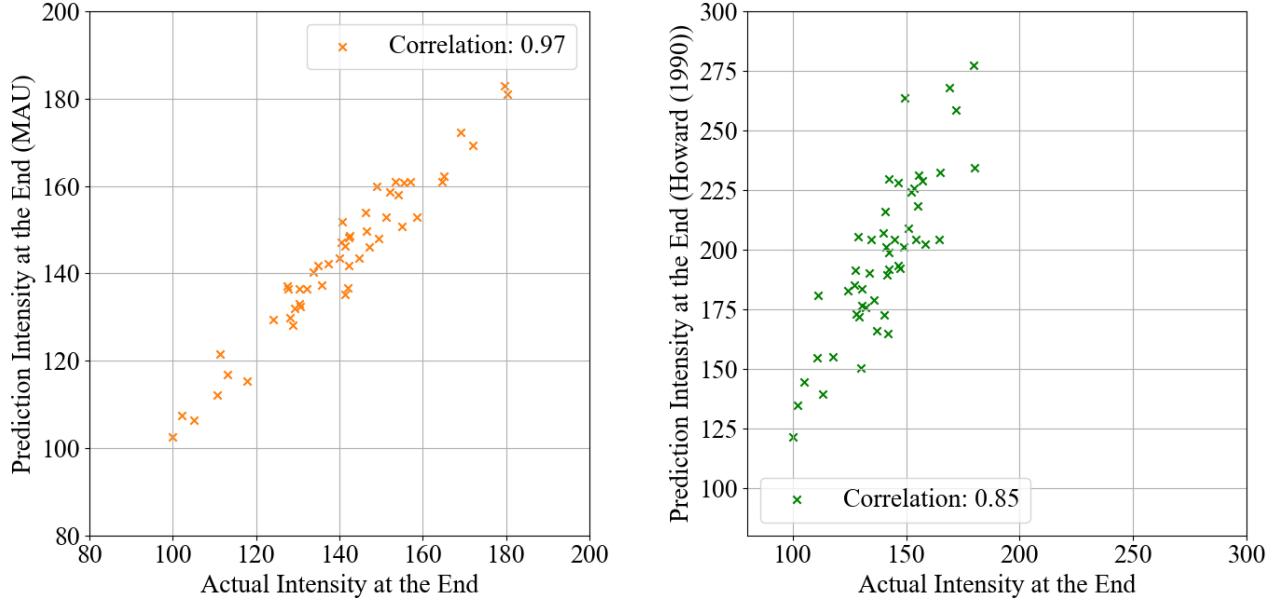


図 4.10: 予測対実測の散布図。縦軸が予測から計算された平均輝度強度、横軸が実際の観測画像から計算された平均輝度強度を表す。

4.4.1.2 画像類似度

画像内での構造的再現度とその時間的变化を評価するために、モデルの出力と対応する時間ステップの実際の観測画像の間の SSIM を計算した。SSIM は、画像の品質評価を目的として、Wang et al. (2004) [2004] で提案された。SSIM は特に構造情報が重要とされる医療画像や衛星画像のような分野で広く使用されている。従来の平均二乗誤差 (MSE) やピーク信号対雑音比 (PSNR) と比較して、SSIM は人間の視覚システムにより近い知覚品質を提供する。従来の手法とは異なり、SSIM は画像の輝度、コントラスト、構造の三つの比較を基にしている。SSIM の定義は以下の通りである:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (4.6)$$

4.4. 実験結果

ここで、 x と y は比較される二つの画像、 μ_x 、 μ_y はそれぞれの画像の平均輝度、 σ_x^2 、 σ_y^2 はそれぞれの分散、 σ_{xy} は共分散である。 C_1 と C_2 は安定性のための小さな定数である。

テストセット全体における、ある時間ステップ t の SSIM の平均を以下のように計算した。

$$\bar{SSIM}_t = \frac{1}{N} \sum_{i=1}^{50} SSIM_{i,t} \quad (4.7)$$

画像類似度も、全球での平均輝度と同様に、全球に対してのみ行い、画像中の背景や外縁部からはみ出すコロナなどはその計算に含まれない。また、平均輝度の場合と同様に、単純差動回転モデルとの比較も同時に行った。このように計算された SSIM の時間推移を図 4.11 に示す。MAU の SSIM、単純差動回転モデルによる SSIM は、共に時間経過とともに単調に減少していくが、MAU による SSIM の方が、最終タイムステップにおいても 0.94 を超えているのに対し、単純差動回転モデルによる SSIM は 0.92 程度である。また、 $t=12$ から 16 の初期段階では、単純差動回転モデルの方が SSIM が高いが、それ以降は逆転し、MAU による SSIM の方が高い。この傾向は平均輝度の場合と概ね同様であり、やはり MAU による予測は画像類似度においてもその低下が緩やかである。

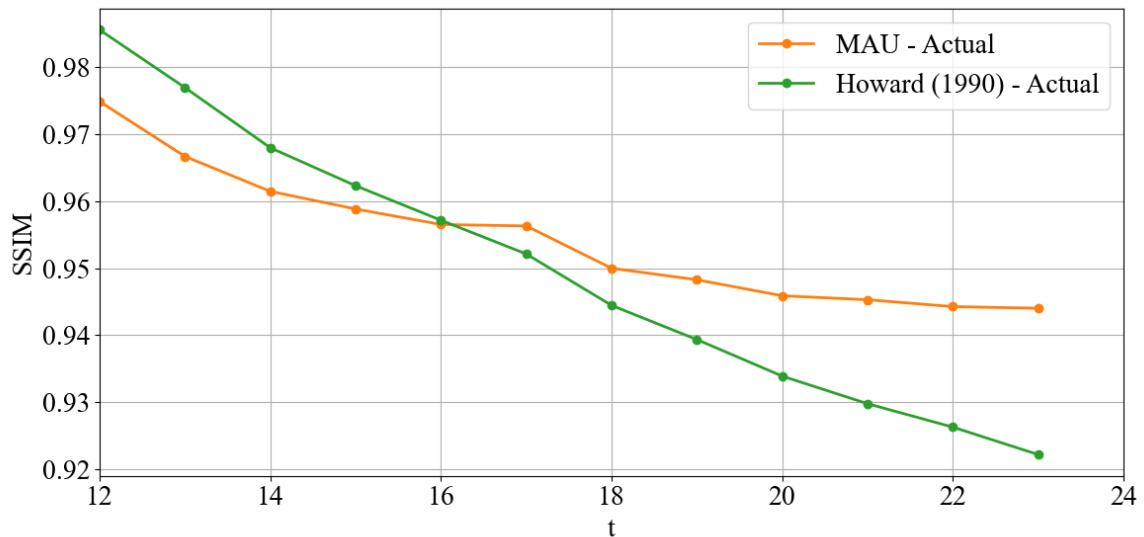


図 4.11: テストセットでの SSIM の時間推移。SSIM は 0 から 1 の値を取り、二つの画像が類似するほど 1 に近づく。横軸が時間ステップ、縦軸が SSIM を表す。

4.4.2 経度依存性の評価

さらに、予測性能が経度ごとにばらつきがあるかを確認するために、経度ごと予測の再現度を評価した。具体的には、Heliographic Stonyhurst 座標系における経度-90°から 90°までの半球を、36°ごとに 5 つのセクターに分割した。分割の概念図を図 4.12 に示す。評価指標には、全球の場合と同様に、平均輝度の平均絶対誤差と、SSIM による画像類似度を用いた。また、それぞれの評価において、単純差動回転モデルとの比較も行った。

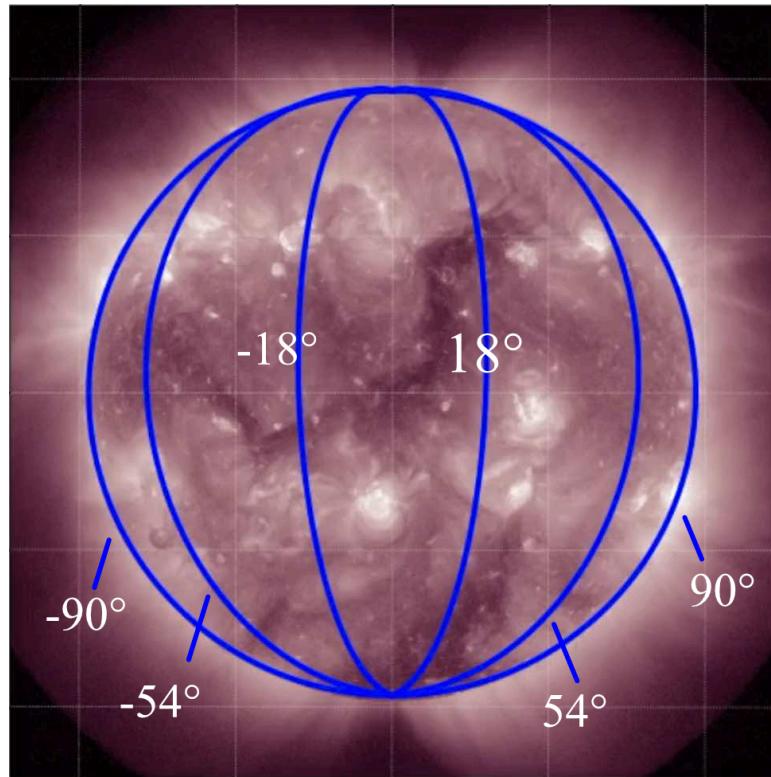


図 4.12: 分割の様子を示した図。Heliographic Stonyhurst 経度-90°から 90°までの半球を、36°ごとに 5 つのセクターに分割した。

4.4.2.1 平均輝度の再現

ここでは、全てのテストセットで各セクターごとの平均輝度を計算し、対応する時間ステップの実際の観測画像との間の絶対誤差を計算した。ここで、ある時間ステップ t 、ある経度セクター l に

4.4. 実験結果

おける平均輝度の絶対誤差 $\bar{E}_{l,t}$ は以下のように定義される:

$$\bar{E}_{l,t} = \frac{1}{N} \sum_{i=1}^{50} |\bar{I}_{\text{Prediction}_{i,l,t}} - \bar{I}_{\text{Actual}_{i,l,t}}| \quad (4.8)$$

$$(4.9)$$

ここで、 i はテストセットのインデックスを表す。また、 $\bar{I}_{\text{Prediction}_{i,l,t}}$ は、テストセット i 、時間ステップ t 、経度セクター l における予測された平均輝度を表し、 $\bar{I}_{\text{Actual}_{i,l,t}}$ は、実際の平均輝度を表す。

このように計算された誤差率の時間推移を図 4.13 に示す。単純差動回転モデルとの比較を行っている。

■経度-90 度から-54 度 -90 度から-54 度のセクターは、東の外縁部（画像に向かって左側）の領域である。この領域に対する誤差の推移を図 4.13a から検証すると、MAU による予測は、最終タイムステップで 10% 程度と、全球平均の誤差率と比べると高くなっている。一方で、単純差動回転モデルによる予測は、最終タイムステップ 80% 程度であり非常に高い。また全タイムステップにおいて、MAU による予測よりも誤差率が高くなっている。

しかし、単純差動回転モデルでは、新しく出現する領域に対するシミュレーションは論理的に不可能であることから、この領域に対するシミュレーション画像は時間の経過によって徐々に欠けていく。そのため、ここでの誤差率を比較する領域は、経度ごとの切り出しからさらに、単純差動回転モデルの出力する領域に合わせて切り取ったものとなっている。特に最終タイムステップ $t=24$ 付近では、その領域は非常に小さくなるため、ここで計算された誤差率は、それ単体で各モデルの性能をを十分に評価できるものではない。そのような理由から、東側外縁部に対する評価は、ここで計算された平均輝度の絶対誤差率のみでは十分に判断できない。

しかし、そのような事情を踏まえても、MAU は差動回転モデルと比較すると顕著な予測能力を示唆する誤差率を示している。この点については、後述する「東側外縁部に対する評価」で詳しく検証する。

■経度-54 度から-18 度 -54 度から-18 度のセクターは、東側の中心部の領域である。この領域は、東側外縁部と比べると、観測される面積が大きいため、より予測は容易になるものの、時間経過によって東側外縁部から移動してくる表面を予測しなければならないため、一定の難しさがある。ここでの誤差の推移を図 4.13b から検証すると、MAU による予測は、最終タイムステップで 5% 程度である。一方で、単純差動回転モデルによる予測は、最終タイムステップで 50% 程度であり、また全タイムステップにおいて、MAU による予測よりも誤差率が高くなっている。このこと

4.4. 実験結果

から、東側外縁部と比べると予測が容易であるこの領域でも、予測性能の時間推移は全体的な傾向と同様であることがわかる。

■経度-18 度から 18 度 -18 度から 18 度のセクターは、太陽の中心部の領域である。この領域は、太陽の中心部であり、観測される面積が最も大きいため、予測は比較的容易であると考えられる。ここで誤差の推移を図 4.13c から検証すると、MAU による予測も単純差動回転モデルによる予測も、全タイムステップで 10% 以下であり、予測難易度は比較的高くないことがわかる。MAU の平均絶対誤差は、最終タイムステップで 6% であり、単純差動回転モデルが約 8.5% である。これから、MAU による予測は、単純差動回転モデルよりも高い精度で予測できていることがわかる。

■経度 18 度から 54 度 18 度から 54 度のセクターは、西側の中心部の領域である。この領域は、東側の中心部と同様に、観測される面積はある程度大きい。また、太陽表面が東から西に移動することから、観測できる面積の大きい太陽中心に近い領域の情報から予測を行うことができるため、東側中心部よりも予測は容易であると考えられる。この領域での平均絶対誤差の推移を図 4.13d から検証する。MAU による予測では、最終タイムステップで、約 6% の誤差、一方、単純差動回転モデルでは約 14% の誤差である。このことから、この領域においても、MAU による予測は単純差動回転モデルよりも高い精度で予測できていることがわかる。

■経度 54 度から 90 度 54 度から 90 度のセクターは、西側の外縁部の領域である。この領域は、東側の外縁部と同様に、実際の面積に対して、観測される面積が非常に小さい。しかし、太陽表面が東から西に移動することから、観測できる面積が大きい領域から小さい領域を予測するため、東側外縁部に比べれば予測は容易であると考えられる。この領域での平均絶対誤差の推移を図 4.13e から検証する。MAU による予測では、最終タイムステップで、約 6% の誤差、一方、単純差動回転モデルでは約 35% の誤差と、性能に大きな差があることがわかる。これは、東側外縁部と同様に、外縁部の領域においては、その予測に大きな不確実性が伴うため、単純なモデルによる決定論的な予測は困難であることを示している。一方で、MAU による予測は、単純差動回転モデルよりも非常に高い精度で予測できていることがわかる。これは、東側外縁部の結果と同様に、MAU による予測は、不確実性の高い領域においても有効であることを示している。

4.4.2.2 画像類似度

全球での場合と同様に、経度ごとにも画像類似度を計算した。その時間推移を図 4.14 に示す。同時に単純差動回転モデルの経度ごとの画像類似度も計算した。全体的な傾向は、全球での場合と同様であり、MAU による予測は、単純差動回転モデルよりも高い精度で予測できていることがわかる。特に、予測難易度の高い東側外縁部 (-90 度から -54 度) および西側外縁部 (54 度から 90 度) に

4.4. 実験結果

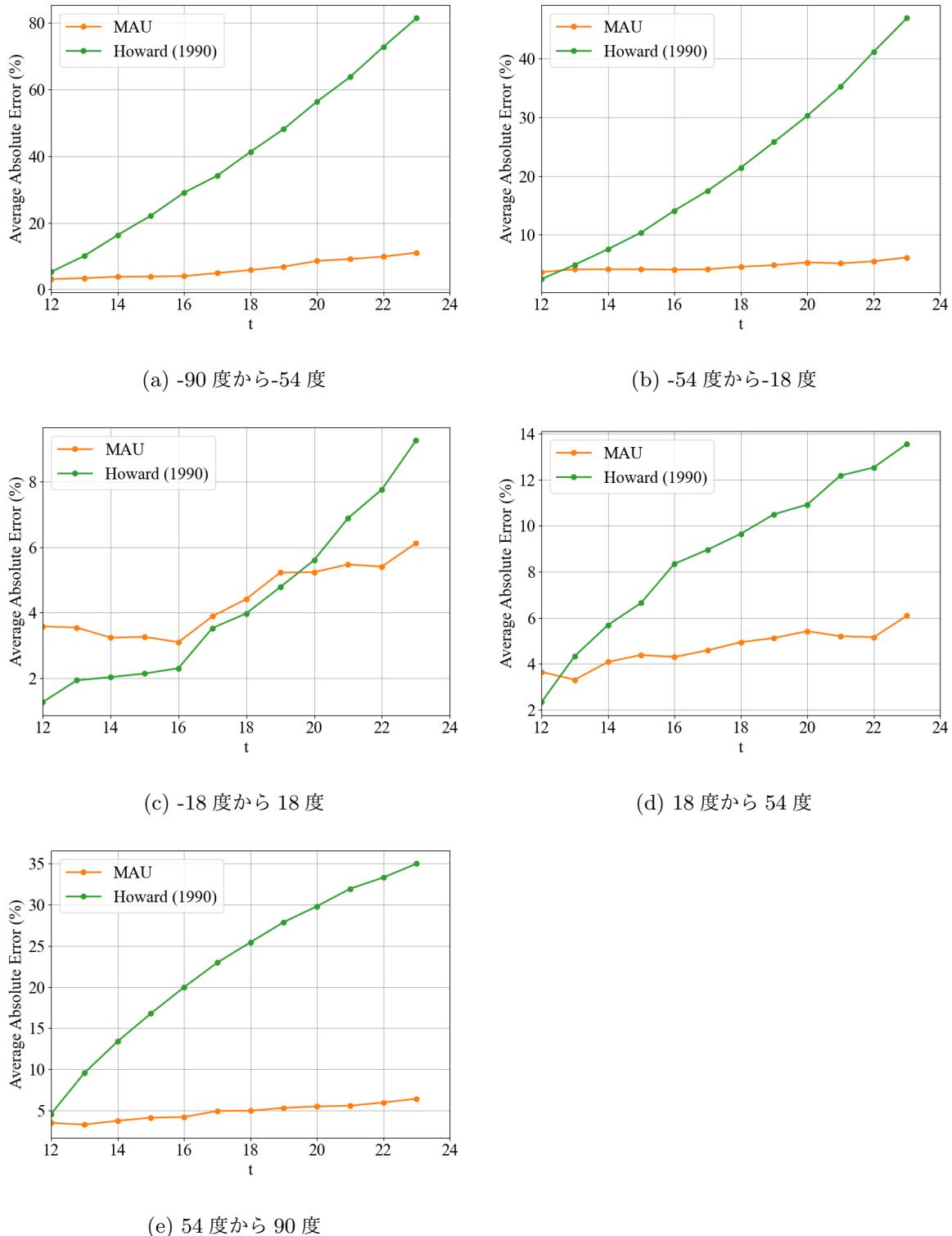


図 4.13: 分割された各セクターにおける平均輝度の絶対誤差の時間推移。横軸が時間ステップ、縦軸が平均絶対誤差を表す。各グラフで縦軸の範囲が異なる。緑線が MAU による予測から計算された絶対誤差、オレンジ線が単純差動回転モデルによるシミュレーションから計算された絶対誤差を表す。

4.4. 実験結果

おいて、MAUによる予測は、単純差動回転モデルの画像類似度に対して顕著な差を示している。

■注 SSIMは、実装の都合上、切り抜いた画像のみではなく、ゼロ埋めされた画像の部分も含めて計算される。そのため、異なる経度セクターでの比較は、画像サイズが異なるため、単純にSSIMを比較することでその経度依存の違いを評価することはできない。同じ経度セクターでの異なるモデルの比較は、画像サイズが同一であるため可能である。

4.4.3 東側外縁部に対する評価

ここまでで、作成した動画予測モデルは、全球での平均輝度や、経度ごとの平均輝度といった定量的な評価において、実際の観測画像を正確に再現できていることを確認した。既存のシンプルなシミュレーションモデルとの比較でも、平均輝度の評価においては、動画予測モデルの優位性を確認できた。

動画予測モデルのシミュレーションモデルに対するさらなる独自の特徴として、望遠鏡の視野に入っていない太陽の球面を生成することができる点が挙げられる。Sunpyによって提供される差動回転シミュレーションモデル *physics.differential_rotation* は、入力された画像の全球面の各ピクセルに対して差動回転を適用することで画像を生成する。そのため、入力時点での望遠鏡の視野に入っていない太陽の球面を生成することができないので、より長い時間スパンでの予測を行うと、東の外縁部から徐々に予測できない領域が広がっていく。また、この領域は、太陽の表面が東から西に向かって回転することから、時間が経過するにつれて、望遠鏡の視野に入る領域が大きくなる領域である。これは、少ない情報から詳細な予測を行う必要があるため、一般的には予測が困難であると考えられる。

これに対して、動画予測モデルは、入力画像の全球面に対して特定の数理モデルを適用するのではなく、過去のデータや全体的な文脈を元に、視野に入っていた領域を含む未来の状態を生成する。さらに、深層学習モデルは、このような不確実性の高いタスクにおいて、決定論的予測を行うシミュレーションモデルよりも有効である場合があることが知られている。ここでは、動画予測モデルが、そのような、入力画像の時点で全球面に見えていない領域に対して予測能力を持つか検証を行うため、生成された画像の東側外縁部から出現する活動領域に対する評価を行う。

4.4.3.1 活動領域に対する視覚的評価

ここでは、東側外縁部から出現する活動領域に注目して、MAUがどのような予測を行っているかを視覚的に評価する。予測の例を図4.15および4.16に示す。ここで示す画像は、上段が実際の観測画像、下段がその予測画像である。出現する活動領域をバウンディングボックスで囲っている。

4.4. 実験結果

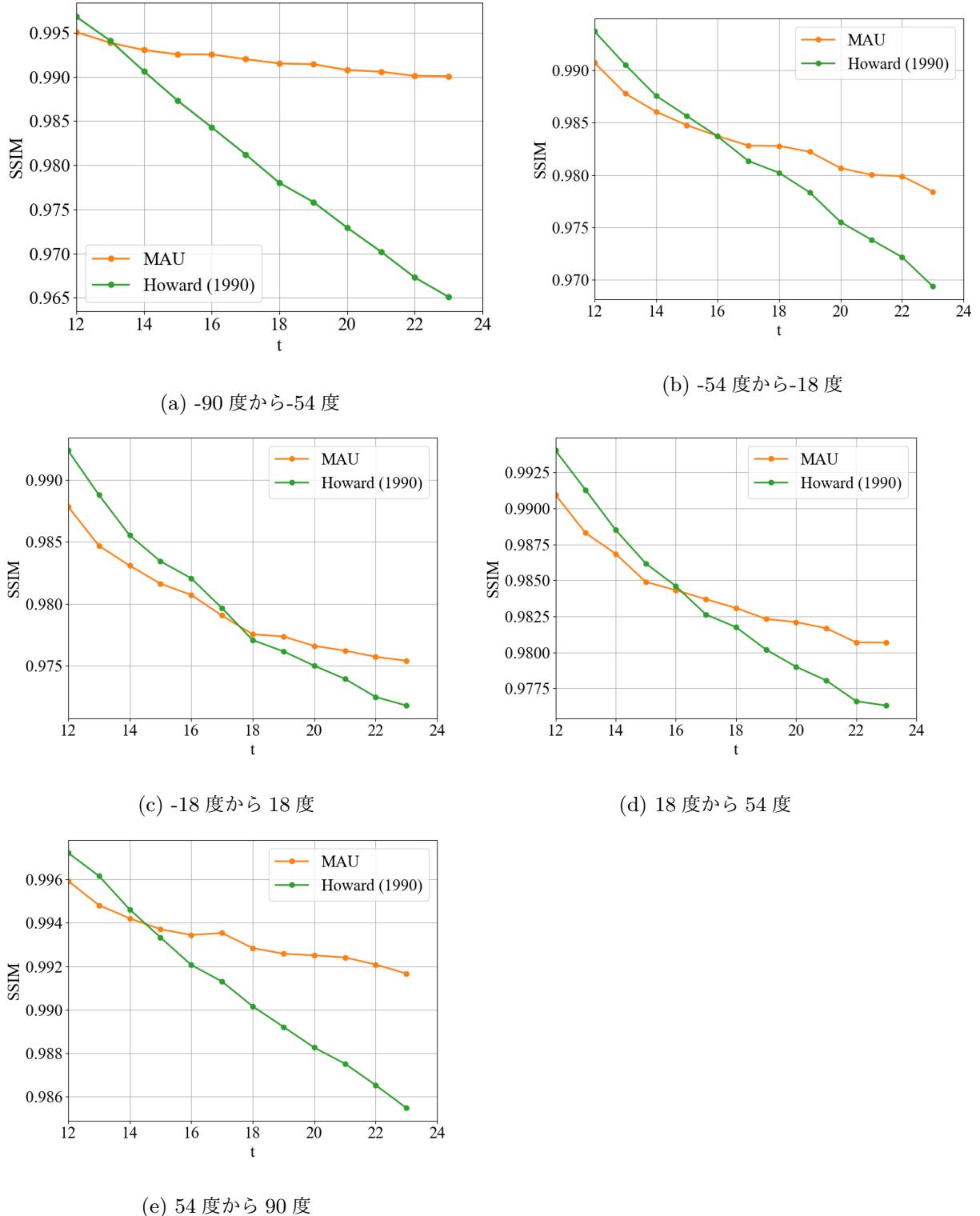


図 4.14: 分割された各セクターにおける SSIM の時間推移。横軸が時間ステップ、縦軸が SSIM を表す。各グラフで縦軸の範囲が異なる。緑線が MAU による予測から計算された SSIM、オレンジ線が単純差動回転モデルによるシミュレーションから計算された SSIM を表す。

4.4. 実験結果

図 4.15 は、活動領域が東側外縁部の北側中緯度帯から出現する例である。図 4.16 は、活動領域が東側外縁部の南側中緯度帯から出現する例である。どちらの例においても、最終入力の時点では全球面には活動領域はほとんど出現していない。しかし、バウンディングボックス内を見ると、その形こそ曖昧になっているものの、活動領域の出現を予測できていることがわかる。この結果から、MAU は、東側外縁部から出現する活動領域に対しても、一定の予測能力を持つことがわかる。

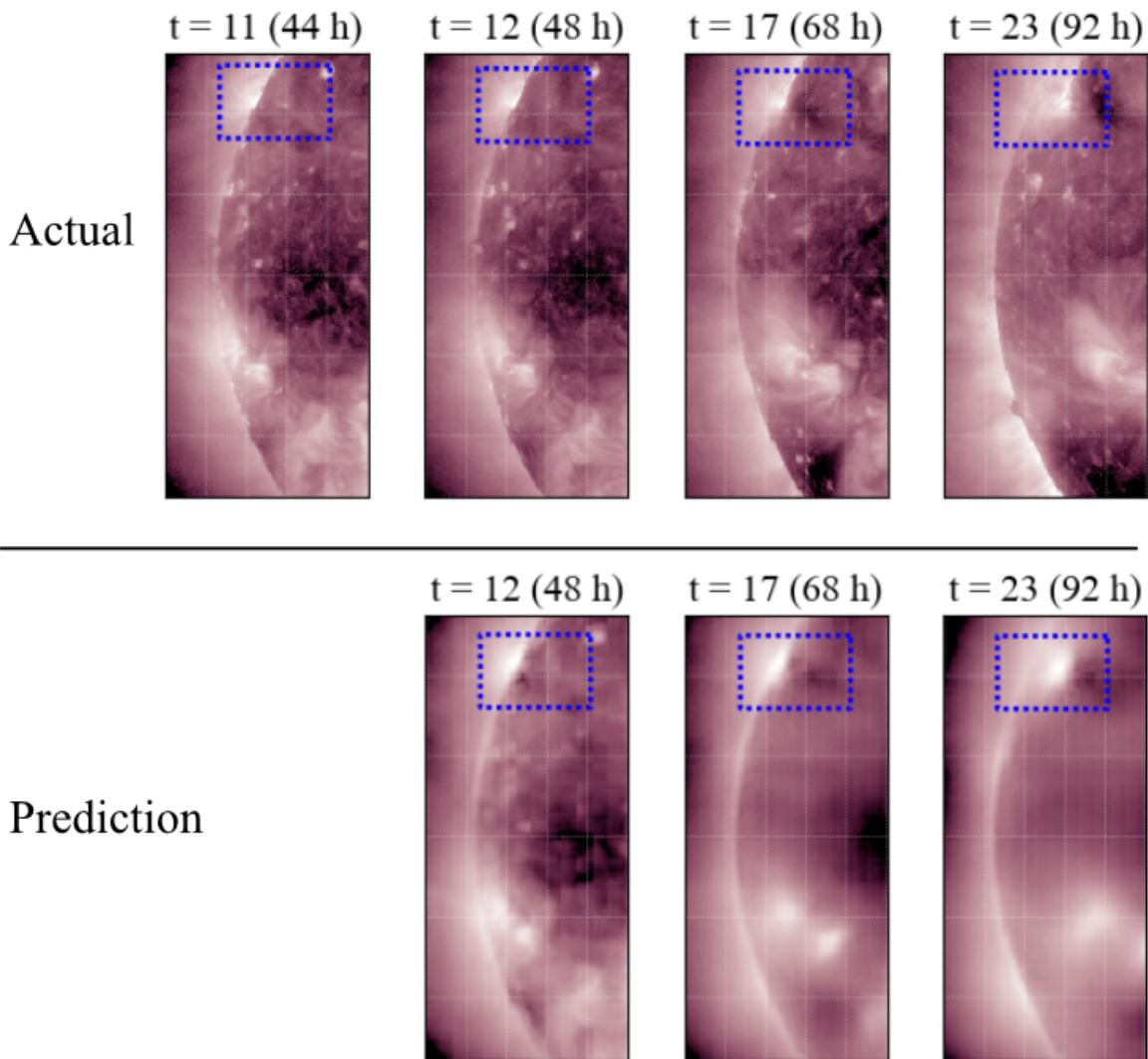


図 4.15: 東側外縁部の北側中緯度帯から出現する活動領域をもつテストセットの例。上段が実際の観測画像、下段がその予測画像である。活動領域を青色破線のバウンディングボックスで囲んでいる。2022 年 11 月 6 日 0 時から 2022 年 11 月 9 日 8 時の期間の画像。

4.4. 実験結果

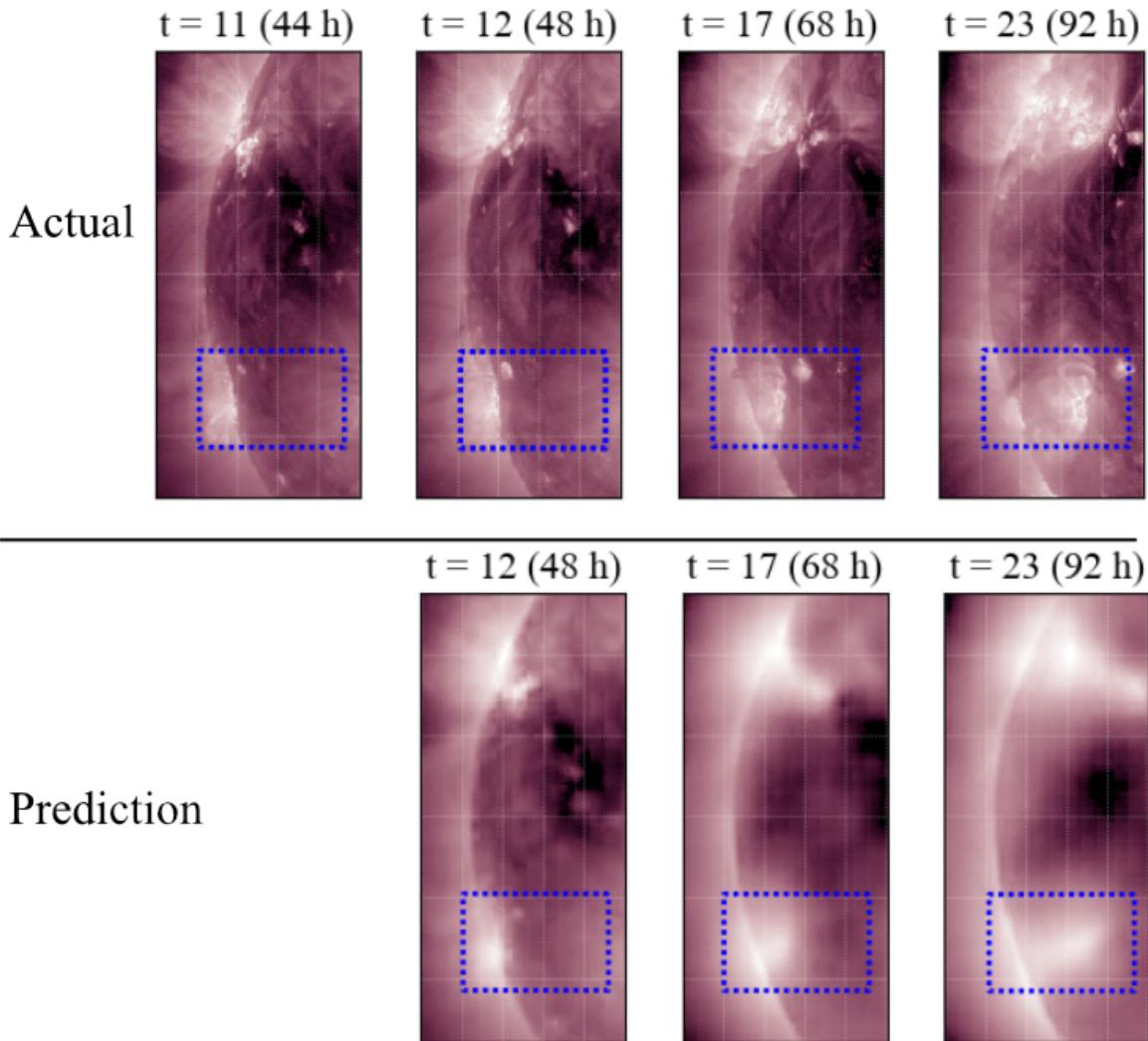


図 4.16: 東側外縁部の南側中緯度帯から出現する活動領域をもつテストセットの例。上段が実際の観測画像、下段がその予測画像である。活動領域を青色破線のバウンディングボックスで囲んでいる。2022 年 12 月 12 日 0 時から 2022 年 12 月 15 日 8 時の期間の画像。

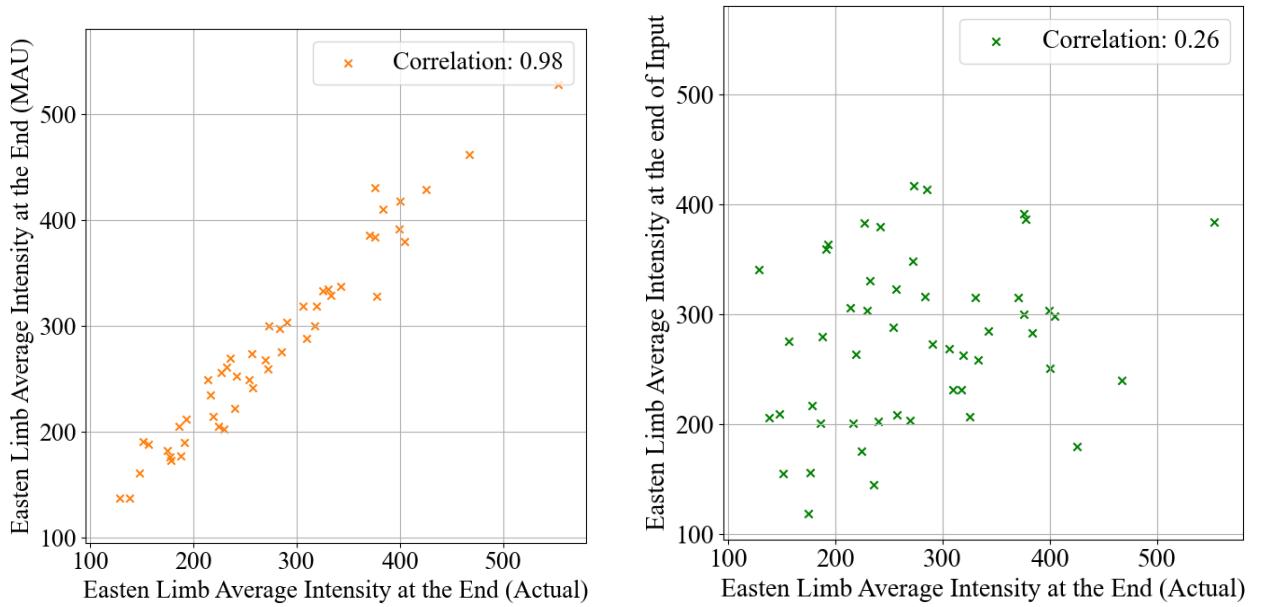
4.4.3.2 予測対実測散布図による定量的評価

さらに、東側外縁部に対する評価を行うために、予測対実測の散布図を作成した。その結果を図 4.17 に示す。左は、MAU による予測画像の東側外縁部の平均輝度と、実際の観測画像の東側外縁部の平均輝度の散布図である。実際の観測画像の東側外縁部の平均輝度と、その 48 時間後の予測画像の東側外縁部の平均輝度は、相関係数 0.98 で、強い相関があることがわかる。また、最終タイムステップにおける実際の観測画像の東側外縁部の平均輝度強度が、その 48 時間の値とどのよう

4.5. 考察

に一致しているかを示す散布図も作成した。その結果を 4.17 右に示す。この散布図から、東側外縁部の平均輝度と、その 48 時間後の東側外縁部の平均輝度は、相関係数 0.26 と相関が弱く、時間経過により容易に変化することがわかる。これは、東側外縁部の平均輝度の変化は、単純なロジックでは予測できないことを示している。

これらの結果から、動画予測モデルは、一定の複雑さを持つ東側外縁部の平均輝度に対しても、高い再現能力を持つことがわかった。



(a) すべてのテストセットの、最終タイムステップでの東側外縁部の平均輝度の予測対実測の散布図。横軸が実際の側外縁部の平均輝度の実測値と、その 48 時間前の実測値の観測画像から計算された平均輝度強度、縦軸が MAU による予測から計算された平均輝度強度を表す。計算された相関係数は 0.98 である。

(b) すべてのテストセットでの、最終タイムステップでの東側外縁部の平均輝度の予測対実測の散布図。横軸が実際の観測画像から計算された平均輝度強度、縦軸が MAU による予測から計算された平均輝度強度を表す。計算された相関係数は 0.26 である。

図 4.17

4.5 考察

ここまで評価結果を表 4.2 にまとめる。さまざまな条件下でのモデルの評価結果から、動画予測モデルの有効性を確認する。

4.5. 考察

表 4.2: 本実験での各評価の結果。MAU は、本研究で使用した動画予測モデルによる予測に対する評価、Howard et al. (1990) は、単純差動回転モデルによるシミュレーションに対する評価を表す。すべての数値は、全テストセットの最終タイムステップでの値を平均したものである。

評価指標	全球	経度ごと				
		-90 to -54	-54 to -18	-18 to 18	18 to 54	54 to 90
平均輝度絶対誤差 ↓						
MAU - 1 波長	3.67	11.0	6.15	6.12	6.09	6.42
Howard et al. (1990)	10.2	81.4	46.9	9.26	13.6	35.0
SSIM ↑						
MAU - 1 波長	0.944	0.99	0.978	0.975	0.981	0.976
Howard et al. (1990)	0.922	0.965	0.969	0.972	0.976	0.985

4.5.1 全球での性能

全球での評価では、平均輝度の絶対誤差と SSIM を計算し、それぞれ単純差動回転モデルのシミュレーション結果と比較を行った。MAU による予測は、全タイムステップで 4% 程度を下回っている。最終タイムステップでの予測対実測の平均輝度の相関係数は 0.98 であり、予測の精度は非常に高いことがわかる。また、単純作動回転モデルによるシミュレーション結果では、最終タイムステップでの平均輝度の絶対誤差は 10% 程度である。

これらの結果から、MAU は、全球の平均輝度の再現に関して有効な能力を持ち、その能力はシンプルなシミュレーションモデルよりも高いことがわかる。これは、モデルが太陽の差動回転に関する再現性能を持ち、太陽の表面の輝度分布の時間変化を学習できていることを示している。

4.5.2 経度に対するロバスト性

経度ごとの評価では、平均輝度の絶対誤差と SSIM を計算し、それぞれ単純差動回転モデルのシミュレーション結果と比較を行った。最終タイムステップでの平均輝度の絶対誤差は、MAU による予測では、-90 度から-54 度で 11.0%、-54 度から-18 度で 6.15%、-18 度から 18 度で 6.12%、18 度から 54 度で 6.09%、54 度から 90 度で 6.42% である。これは、単純差動回転モデルと比較すると、全球での場合よりも大きい差がついている。また、単純差動回転モデルは、両端の外縁部の領域では誤差が非常に大きく、中心の領域に近づくほど誤差が小さくなっている。しかし、MAU による予測では、東の外縁部では 11.0% と大きくなっているもの、他の領域では 6% 程度であり、大きな精度の偏りは見られない。このことから、MAU は、経度に対してロバストな予測能力を持つ

4.5. 考察

ことがわかる。

太陽活動は、太陽の全域にわたって多様な形で発生し、地球への影響もその活動の種類や位置によって異なるため、全経度にわたる精度の高い予測は宇宙天気予報において重要である。MAU が示した経度に対するロバスト性は、太陽の動的な挙動や複雑な活動パターンを理解する上で大きな利点となる。

4.5.3 東側外縁部における予測性能

東側外縁部 (-90 度から -54 度) の最終タイムステップでの平均輝度の予測対実測の相関係数は 0.98 であり、高い相関を示した。また、視覚的評価において、東側外縁部から出現する活動領域に対して一定の予測能力を持つことがわかった。

ベースラインとして計算した、入力シーケンスの最終タイムステップでの実際の観測データの平均輝度と、その 48 時間後の実際の観測データの平均輝度の相関係数は 0.26 であった。これは、東側外縁部の平均輝度は時間経過によって容易に変化することを示しており、その予測の難しさを示唆している。このような困難性は、以下のような理由によるものと考えられる。

■新しい観測面の出現 太陽が西から東に自転することから、東側外縁部は常に新しい太陽表面が観測範囲に入ってくる領域である。予測モデルにおいては、その直前の観測データを用いて予測を行う。そのことから、そのデータが非常に限られており、入力データに存在しない観測面を予測しなければならない、という点は、予測の難易度を直接的に高め、その精度を低下させる可能性がある。

■面積の小ささと回転による増加 東側外縁部は、太陽の表面が望遠鏡の視野に入る角度が急になる領域である。そのため、実際の太陽表面の面積に対して、観測される面積が相対的に非常に小さくなる。その上、角度は徐々に緩やかになるため、その未来の状態を予測する場合、その少ない情報から詳細な予測を行う必要がある。

■角度の解釈 太陽の中心部では、その面が視線に対して直角に近くなるため、活動領域などの太陽表面の状態が、観測者にとって最もわかりやすい形で観測される。一方で、東側外縁部では、その面が視線に対して鋭角になるため、太陽表面の状態が実際の太陽表面の状態とは異なる形で観測される。このため、東側外縁部で得られる観測データから実際の状態を正確に推定するには、観測角度による歪みを考慮した、角度の意味的解釈が必要になる。これは、太陽の中央部に比べて、観測データの解析がより複雑になることを意味しており、予測モデルにとっても大きな課題となる。

4.5. 考察

このような理由から、東側外縁部の予測は、その他の領域の予測よりも難しいと考えられる。しかし、上述のように、この領域においても MAU は良好な予測性能を示した。当然他の経度分割領域と比較すると相対的には精度は劣るもの、この性能は当初の予測とは異なる結果であり、注目すべき点である。

この結果について、以下のような理由が考えられる。

■高度な学習能力による予測の最適化 MAU が示した東側外縁部における予測精度は、モデルが高度な学習能力を持ち、限られた観測データからの一般化を行えることを示している。このような、複雑で不確実性が高く、データの少ない状況においても、未知の状態の推定が可能であることは、宇宙天気予報において重要な能力である。

■外縁部を越える高温プラズマからの予測 MAU は、図 4.18 に示すように、全球の範囲外から外縁部を超えて観測される高温プラズマの輝きを利用して、通常は観測できない全局面の裏側に隠れた活動領域の存在を予測している可能性がある。太陽コロナは、太陽表面の活動状態に密接に関連しており、特に活動領域周辺では明るく輝くことが知られている。MAU がこの輝きと活動領域の出現のパターンを学習し、それを用いて観測されていない活動領域の位置を予測していると考えられる。このように、MAU は単なる表面の画像だけでなく、高温プラズマの輝きといった間接的な情報をも活用し、より正確な太陽活動の予測を行っている可能性がある。

これらの点は、MAU が差動回転などの比較的単純な現象だけではなく、太陽の物理的な特性やダイナミクスを理解し、それを基に予測を行っていることを示唆している。

4.5.4 複雑なシステムに対する確率的予測の有効性

ここでは、MAU の時間経過に対するロバスト性から、複雑なシステムを持つ太陽の表面活動に対して、確率的予測を行う深層学習モデルが有効である可能性を考察する。

MAU による予測画像から計算された平均輝度の絶対誤差を示した図 4.9 や SSIM を示した図 4.11 から確認すると、MAU は、時間が経過しても精度低下の傾向が単純差動回転モデルと比較して緩やかであることがわかる。また、出力の初期段階では単純差動回転モデルの方が誤差の小さい予測を行っている場合にも、時間経過により MAU の予測精度が単純差動回転モデルの予測を上回っている。このような結果は、我々の深層学習を用いた動画予測モデルが時間経過に対してロバストな予測能力を持つことを示している。

太陽の表面活動は非常に複雑で変動が激しいことで知られている。このようなシステムを物理的シミュレーションによってモデル化を行う研究は、????や????などを代表に盛んに行われている。こ

4.5. 考察

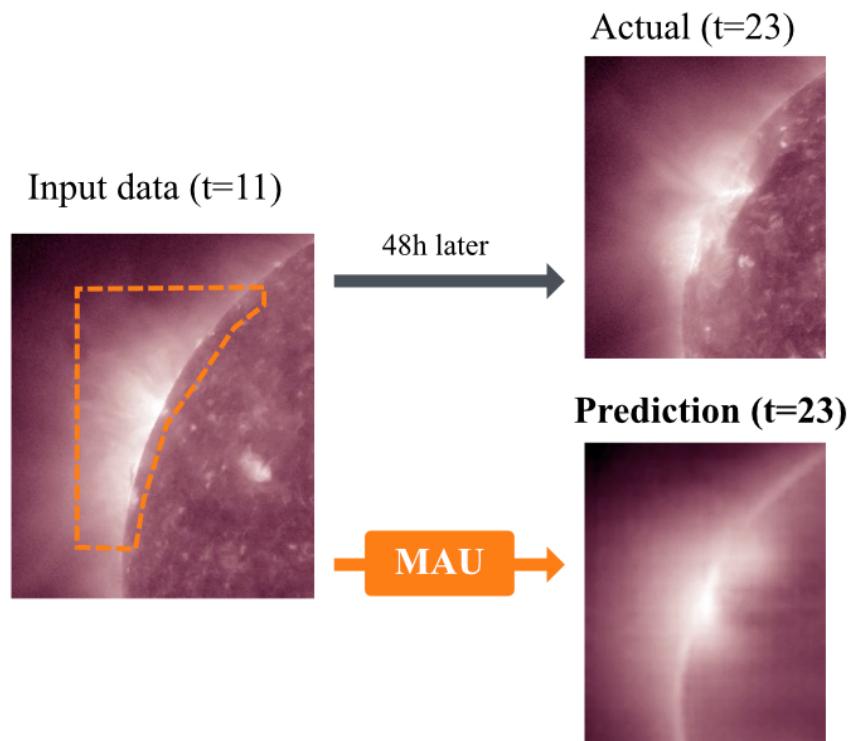


図 4.18

これらを含む多くの物理モデルでは、特定の物理的な仮定や近似に基づいて決定論的な予測を行う。しかし、その複雑性やカオス的性質から、太陽全球といった広範囲にわたり完全に正確なシミュレーションを行うことは難しい。また、そのシミュレーションは非常に高コストであるのが一般的である。そのような性質により、定式化が十分に正確でない場合や、限られた計算リソースのもとで実行する場合、そのモデルは大きな誤差を含む予測を行ってしまう恐れがある。

この状況において、MAU のような確率的予測を用いる深層学習モデルが、太陽活動の予測において新たな視点を提供する。MAU は、一定のぼやけを含む予測を行う。これは不確実性の表現として重要である。ぼやけを含むことで、視覚的なリアルさは一定程度損なわれるが、不確実性の高い領域や、長期的な時間範囲において、より実際に近い物理的数値を含む予測を行うことができる。物理的シミュレーションを用いるモデルでは、決定論的予測を行うため、ぼやけを含まないが、それは必ずしも正確な予測を行うことを意味しない。動画予測モデルは、長期間にわたって視覚的にリアルで高精細な予測を行うことは難しいが、その代わりに、不確実性を考慮した予測を行うことができる。

また、本研究での MAU は、その学習の完了に 10 時間単位の時間と高性能な GPU を要求するが、学習済みモデルがテストデータに対して予測を行うのに要する時間は 1 秒未満である。これは、シミュレーションにスーパーコンピュータベルの計算リソースを要する物理モデルと比較すると非常

4.5. 考察

に高速であり低コストである。

深層学習を用いた動画予測モデルのこのような特徴は、太陽活動のような複雑なシステムに予測に対して有効であると考えられる。

第 5 章

Motion-Aware Unit を用いた 3 波長を入力とした紫外線像の全球時系列予測

5.1 実験概要

ここでは、 171 \AA 、 193 \AA フィルターで得られたデータを追加で利用し、3 波長の入力データから 211 \AA の波長データに対する予測を行った。これらの波長は、太陽のコロナ領域における異なる温度帯を観測するためのものであり、予測モデルに多様な物理的情報を提供することが期待される。 171 \AA の波長は、太陽のコロナにおける温度が約 60 万 K の領域を捉えるのに特化しており、 193 \AA の波長は約 100 万 K の領域を捉える。これらの波長から得られる情報を組み合わせることにより、単一の波長では捉えられない層間の相互作用を捉え、より高い精度での予測を可能にすることを期待する。

モデルには先の実験と同じく、MAU を用いる。入力は 3 波長、すなわち画像的には 3 チャンネルである。目的となる出力は 211 \AA の波長のみであるが、MAU は 3 チャンネルを出力する。これは、「出力シークエンスのタイムステップ 1 以降では、直前のモデルの出力を入力データとして扱う」という動画予測モデルの一般的な性質によるものである。このような性質から、 211 \AA の波長のみを出力として扱うために、出力された 3 チャンネルのうち、 211 \AA の波長のみを抽出するという処理を行った図 5.1 に本実験の概念図を示す。

5.2 実験設定

各ハイパーパラメータの設定を表 5.1 に示す。チャンネル数は入力波長に合わせて 3 である。バッチサイズは 2 に変更している。これは、NVIDIA RTX A6000 のメモリ容量の制約によるものである。データに関しても、データ数の増減による影響がないように、前回の実験と同じ期間のデータ

5.2. 実験設定

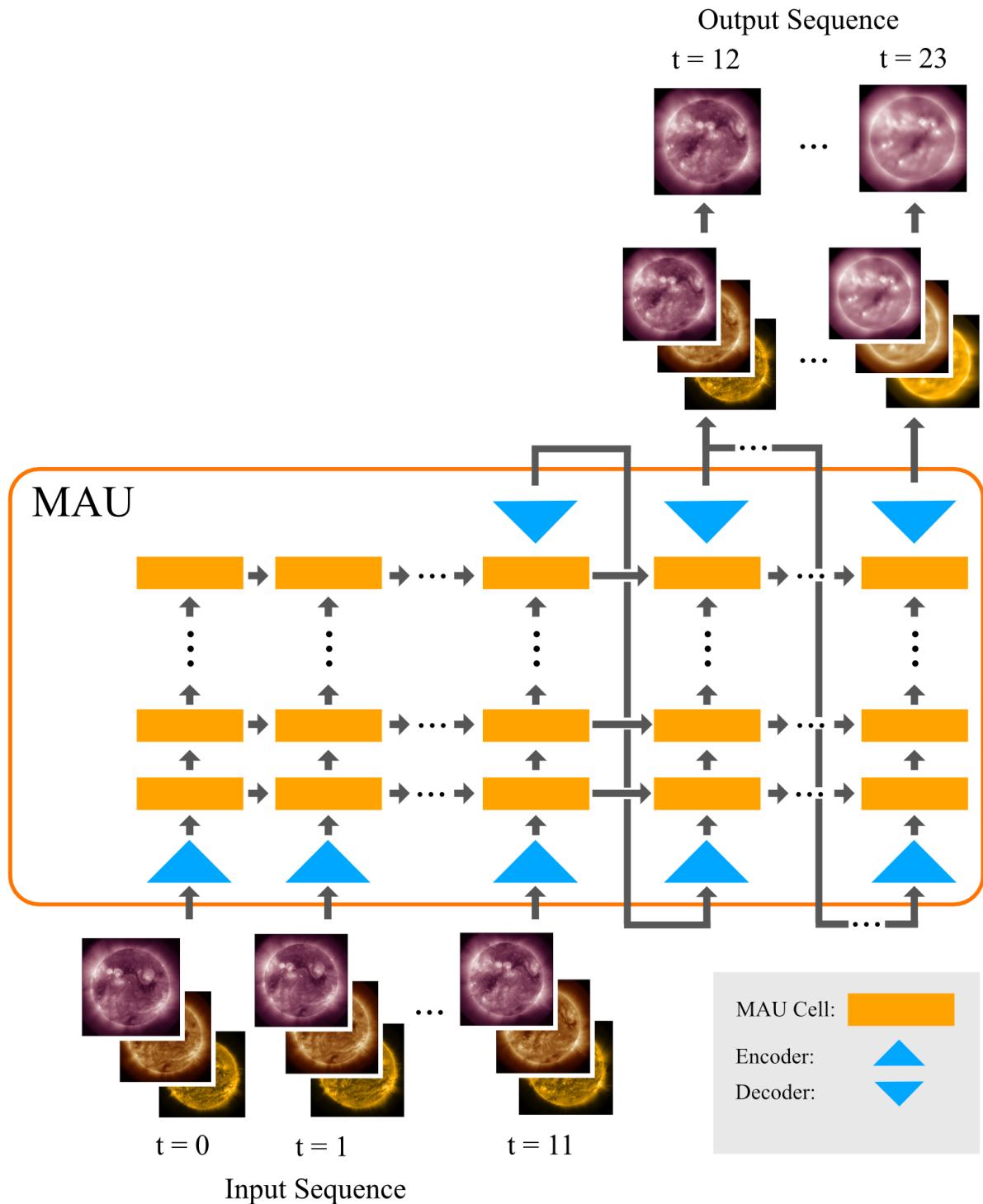


図 5.1: 本実験の概念図。3 波長の観測画像を入力として、211 Å の波長の観測画像を予測する。

5.3. 学習の推移

ハイパーパラメータ	値
バッチサイズ	2
エポック数	100
学習率	0.0005
損失関数	MSE
チャンネル	3
カーネルサイズ	(5, 5)
MAU Cell 数	16

表 5.1: 本実験でのハイパーパラメータ設定。基本的には前実験と同様であるが、チャンネル数が 1 から 3 に変更されている。

タを用いた。欠損期間なども同様である。

5.3 学習の推移

学習は図 5.2 のように推移した。学習の初期段階では、学習データに対する損失関数の値が急激に減少しているが、学習が進むにつれて収束に向かって緩やかに減少していることがわかる。また、学習損失、検証損失ともに、安定的に減少していることがわかる。学習には NVIDIA RTX A6000 を用い、完了までに約 60 時間を要した。

5.4 実験結果

図 5.3 および図 5.4 に、この実験での出力例を示す。

モデルの出力は、視覚的には実際の観測画像と概ね合致している。この実験における評価では、前回の実験と同様の評価を行った。

5.4.1 全球での評価

はじめに全球での評価を行った。前回実験と同様に、まず輝度強度の平均値と実際の平均値との誤差、SSIM を計算した。さらに単純差動回転モデルとの比較も行った。また、これらの値の時間経過に対する変化を観察し、より不確定性の高い将来の予測に対しても動画予測モデルが有効であるかを検証した。

5.4. 実験結果

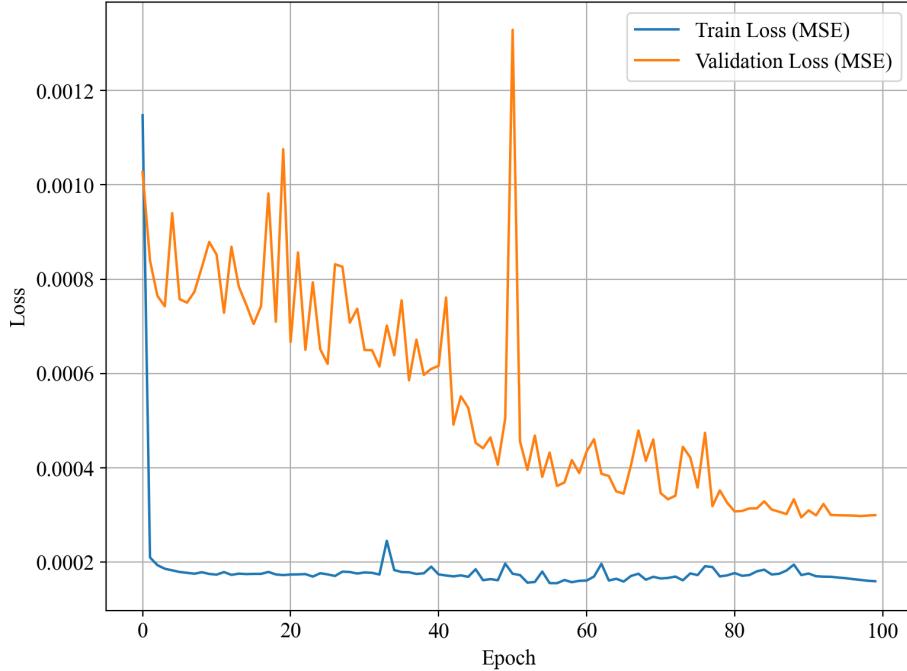


図 5.2: 本実験での、学習データ、検証データでの損失関数の推移。どちらも安定的に減少している。

5.4.1.1 平均輝度の再現

モデルの出力の全球での平均輝度と、実際の観測画像との絶対誤差の推移を図 5.5 に示す。これは、50 のテストセットに対して、各テストセットに含まれる各画像の全球での平均輝度を計算し、その時間ステップごとの平均値を取ったものである。また、前回実験と同様に、モデルの予測性能をさらに詳細に評価するために、シンプルな差動回転モデルとの比較を行った。

さらに、出力シークエンスの最後のタイムステップにおいて、単純差動回転モデルによるシミュレーションと、実際の観測画像との差異を観察し、動画予測モデルによる出力と比較した。このタイムステップは、出力の最後のタイムステップであり、最も不確定性の高い予測である。その散布図 5.6 に示す。

5.4.1.2 画像類似度

前回実験と同様に、画像内での構造的再現度とその時間的変化を評価するために、モデルの出力と対応する時間ステップの実際の観測画像の間の SSIM を計算した。SSIM の推移を図 5.7 に示す。画像類似度は、全球での平均輝度と同様に、全球に対してのみ行い、画像中の背景や外縁部からは

5.4. 実験結果

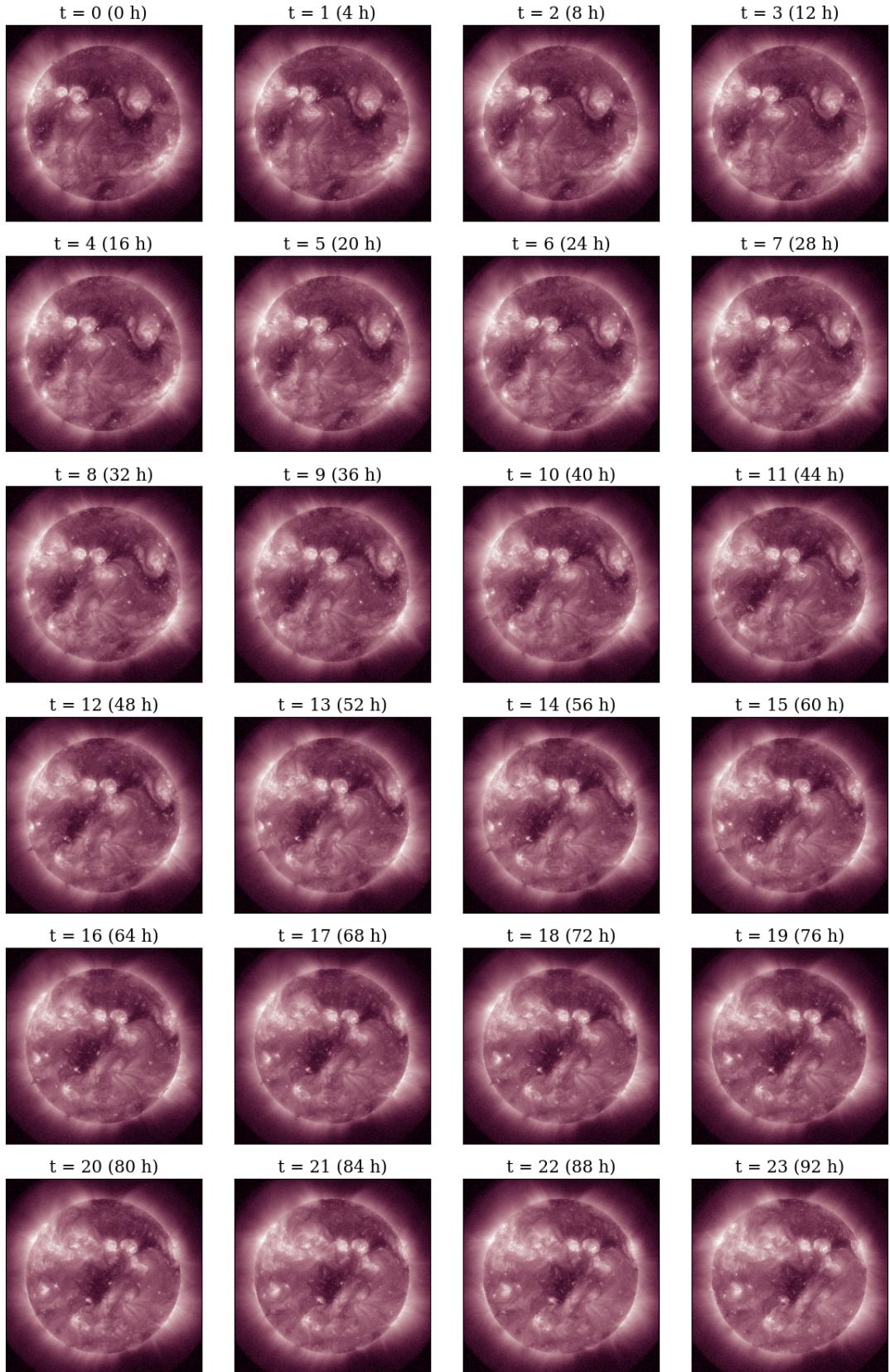


図 5.3: 実際の観測画像の例。2022 年 2 月 18 日 0 時から 2022 年 2 月 22 日 20 時の期間から 4 時間毎にサンプリングされている。この $t=0$ から $t=11$ までをモデルに入力データとして渡している。モデルはその入力データを元に、 $t=12$ から $t=23$ の 12 枚の画像を予測する。 $t=12$ 以降の実際の観測画像はモデルに渡されない。

5.4. 実験結果

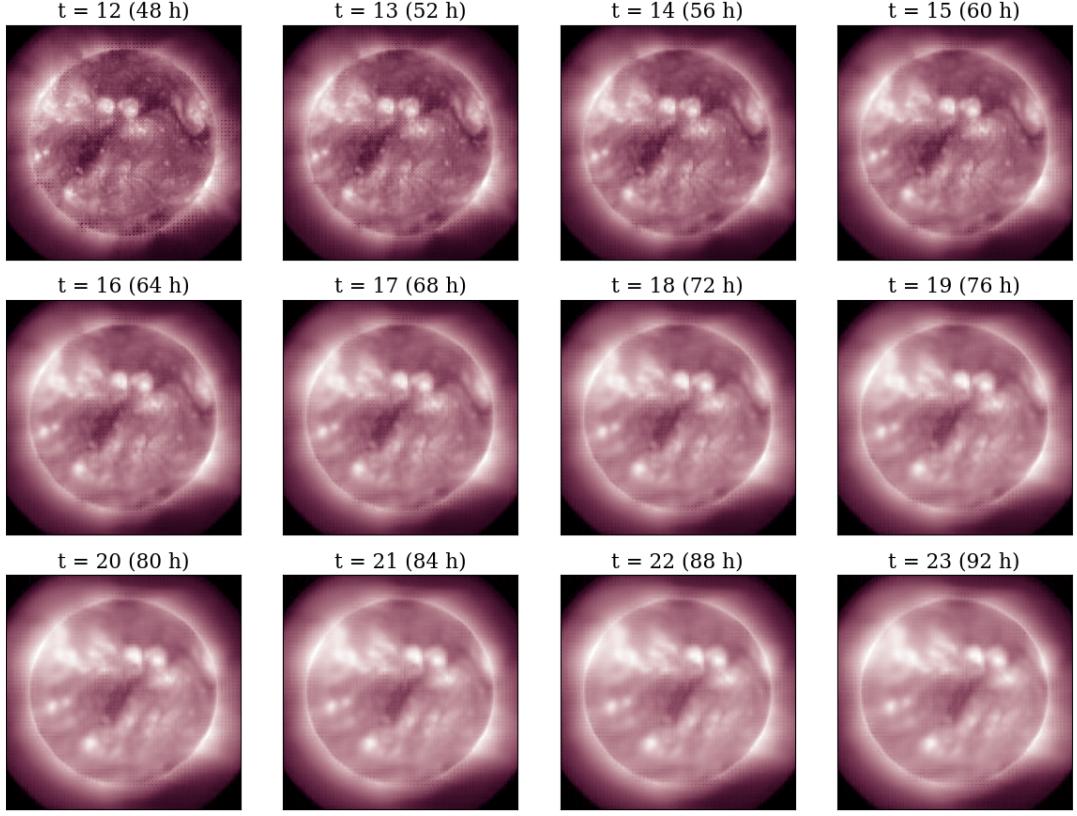


図 5.4: MAU による予測画像。対応するタイムステップ t の観測画像 (図 5.3) と比較することでモデルの再現度を視覚的に評価することができる。大規模な構造は概ね実際の観測画像と合致している。モデルの特性により、時間経過とともに少しずつ予測が不安定になり、ぼやけた見た目になる。

み出すコロナなどはその計算に含まれない。

5.4.2 経度依存性の評価

前回実験と同じく、予測性能が経度ごとにばらつきがあるかを確認するために、経度ごと予測の再現度を評価した。分割の方法は前回実験と同様である。評価指標には、平均輝度の誤差と、その単純差動回転モデルとの比較を用いた。

5.4.2.1 平均輝度の再現

ここでは、全てのテストセットで各セクターごとの平均輝度を計算し、対応する時間ステップの実際の観測画像との絶対誤差を計算した。誤差率の時間推移を図 5.8 に示す。同時に単純差動回転モ

5.4. 実験結果

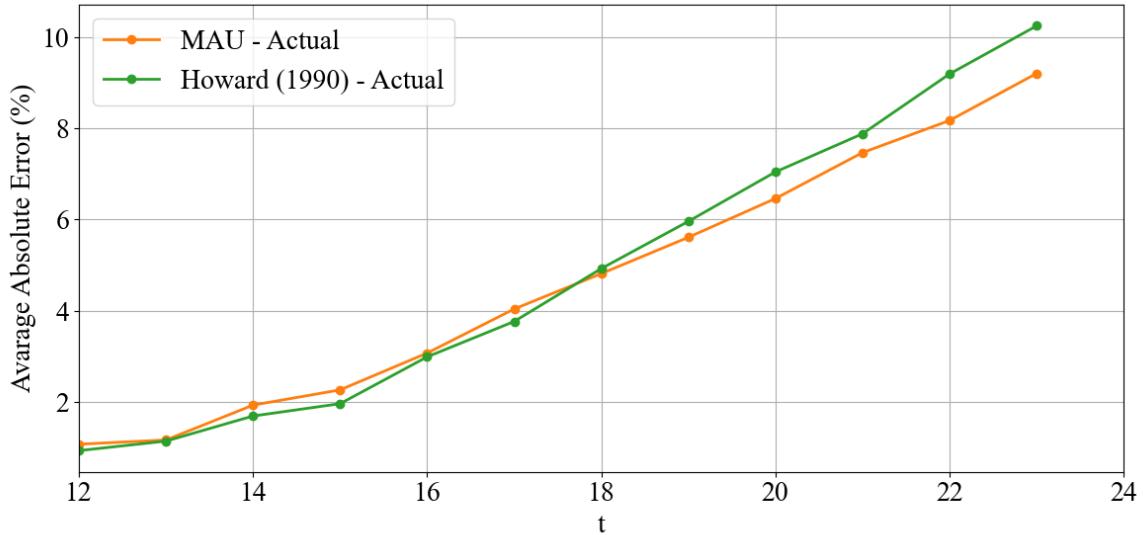
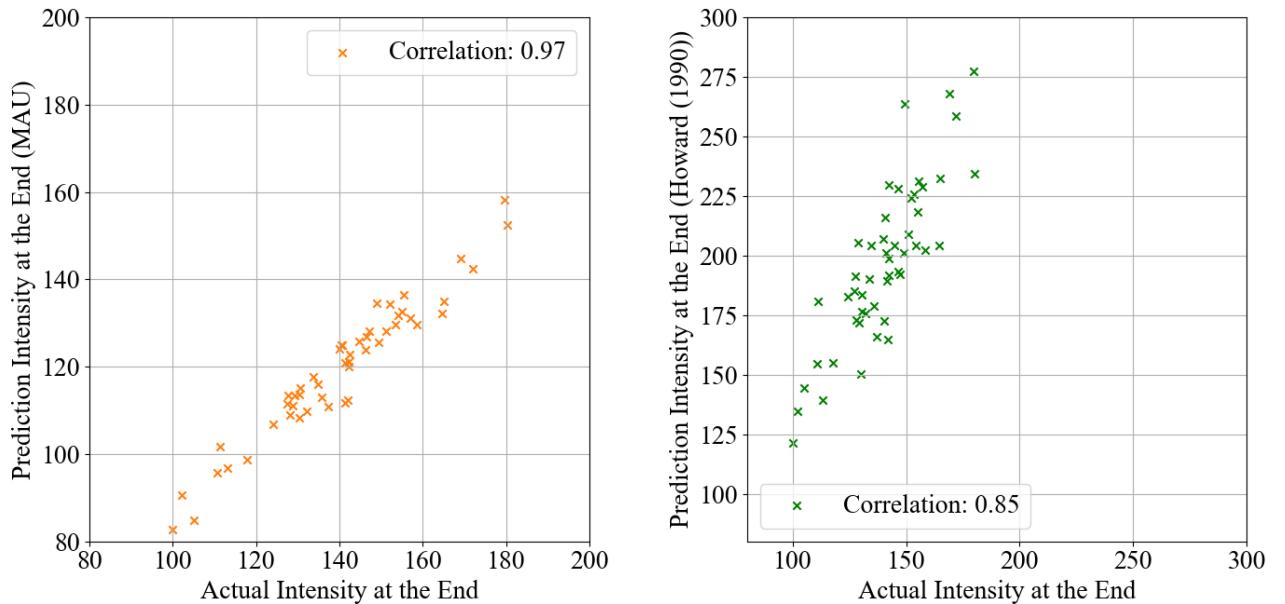


図 5.5: MAU によるテストセットの予測画像と実際の観測画像の平均絶対誤差 (オレンジ) と、単純差動回転モデルと実際の観測画像の平均絶対誤差 (緑)。



(a) MAU による、テストセットの最終ステップにおける全平均輝度の予測対実測の散布図。計算された相関係数は 0.97 である。
(b) 単純差動回転モデルによる、テストセットの最終ステップにおける全平均輝度の予測対実測の散布図。計算された相関係数は 0.85 である。

図 5.6: 予測対実測の散布図。縦軸が予測から計算された平均輝度強度、横軸が実際の観測画像から計算された平均輝度強度を表す。

5.4. 実験結果

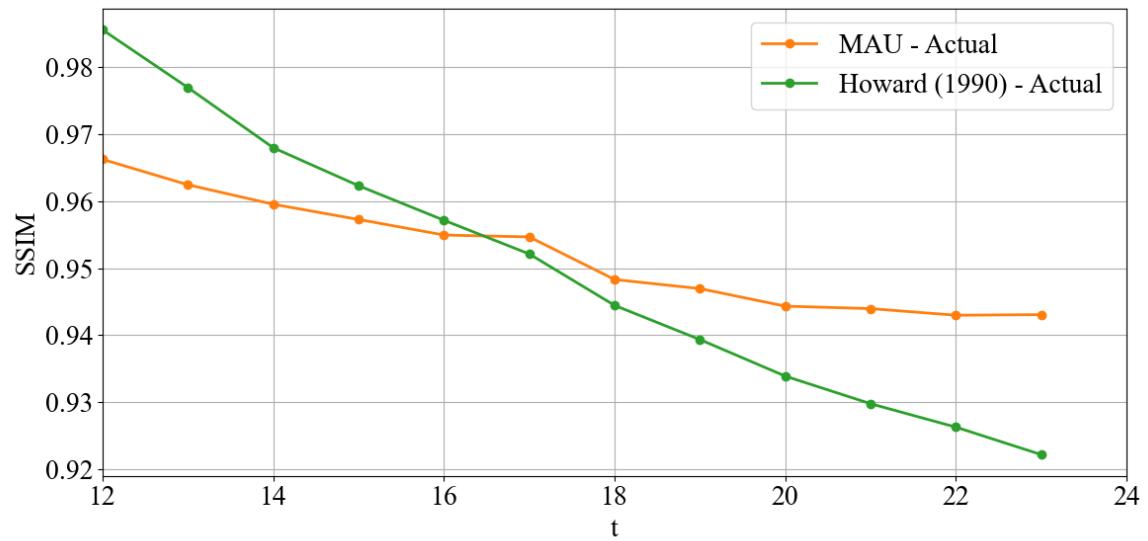


図 5.7: テストセットでの SSIM の時間推移。SSIM は 0 から 1 の値を取り、二つの画像が類似するほど 1 に近づく。横軸が時間ステップ、縦軸が SSIM を表す。

デルの経度ごとの画像類似度も計算した。

5.4.2.2 画像類似度

全球での場合と同様に、経度ごとにも画像類似度を計算した。その時間推移を図 5.9 に示す。同時に単純差動回転モデルの経度ごとの画像類似度も計算した。

5.4.3 東側外縁部に対する評価

5.4.3.1 視覚的評価

ここでは、動画予測モデルが、東側外縁部から出現する活動領域に対して、どのような予測を行っているかを視覚的に評価する。いくつかの例を図 5.10 および 5.11 に示す。

5.4.3.2 散布図

さらに、東側外縁部に対する評価を行うために、予測対実測の散布図を作成した。その結果を図 5.12 に示す。左は、MAU による予測画像の東側外縁部の平均輝度と、実際の観測画像の東側外縁部の平均輝度の散布図である。

5.4. 実験結果

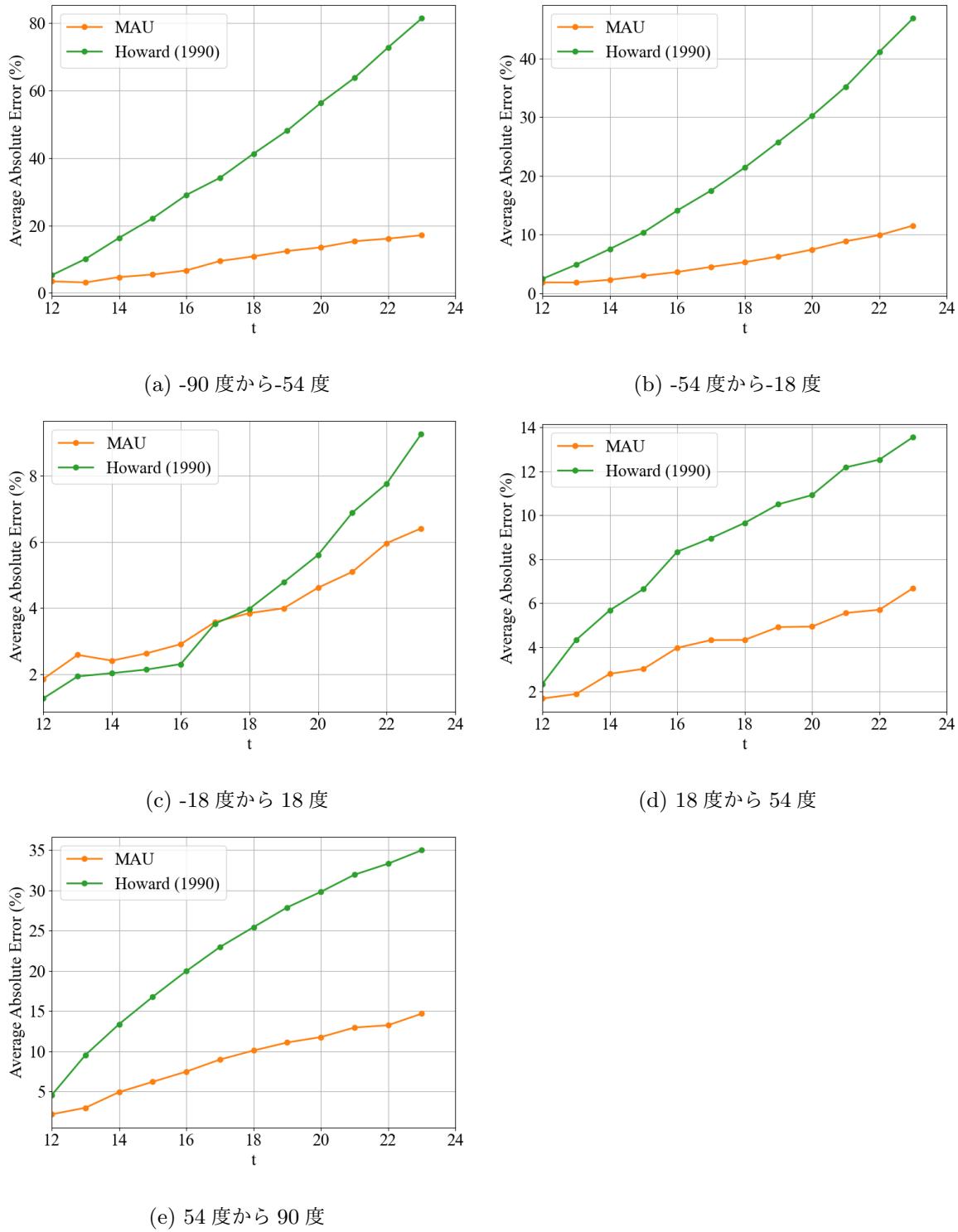


図 5.8: 分割された各セクターにおける平均輝度の絶対誤差の時間推移。横軸が時間ステップ、縦軸が平均絶対誤差を表す。各グラフで縦軸の範囲が異なる。緑線が MAU による予測から計算された絶対誤差、オレンジ線が単純差動回転モデルによるシミュレーションから計算された絶対誤差を表す。

5.4. 実験結果

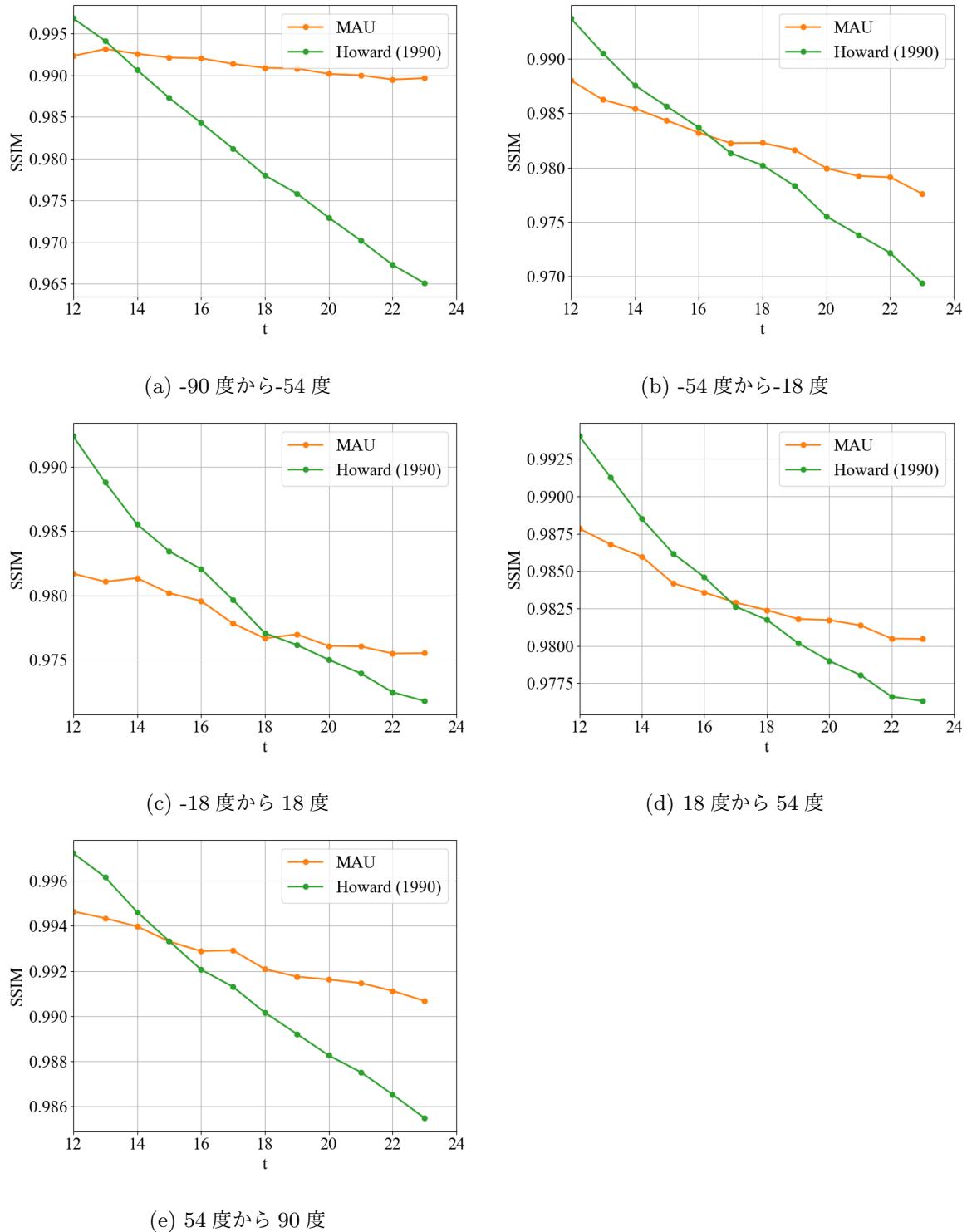


図 5.9: 分割された各セクターにおける SSIM の時間推移。横軸が時間ステップ、縦軸が SSIM を表す。各グラフで縦軸の範囲が異なる。緑線が MAU による予測から計算された SSIM、オレンジ線が単純差動回転モデルによるシミュレーションから計算された SSIM を表す。

5.5. 考察

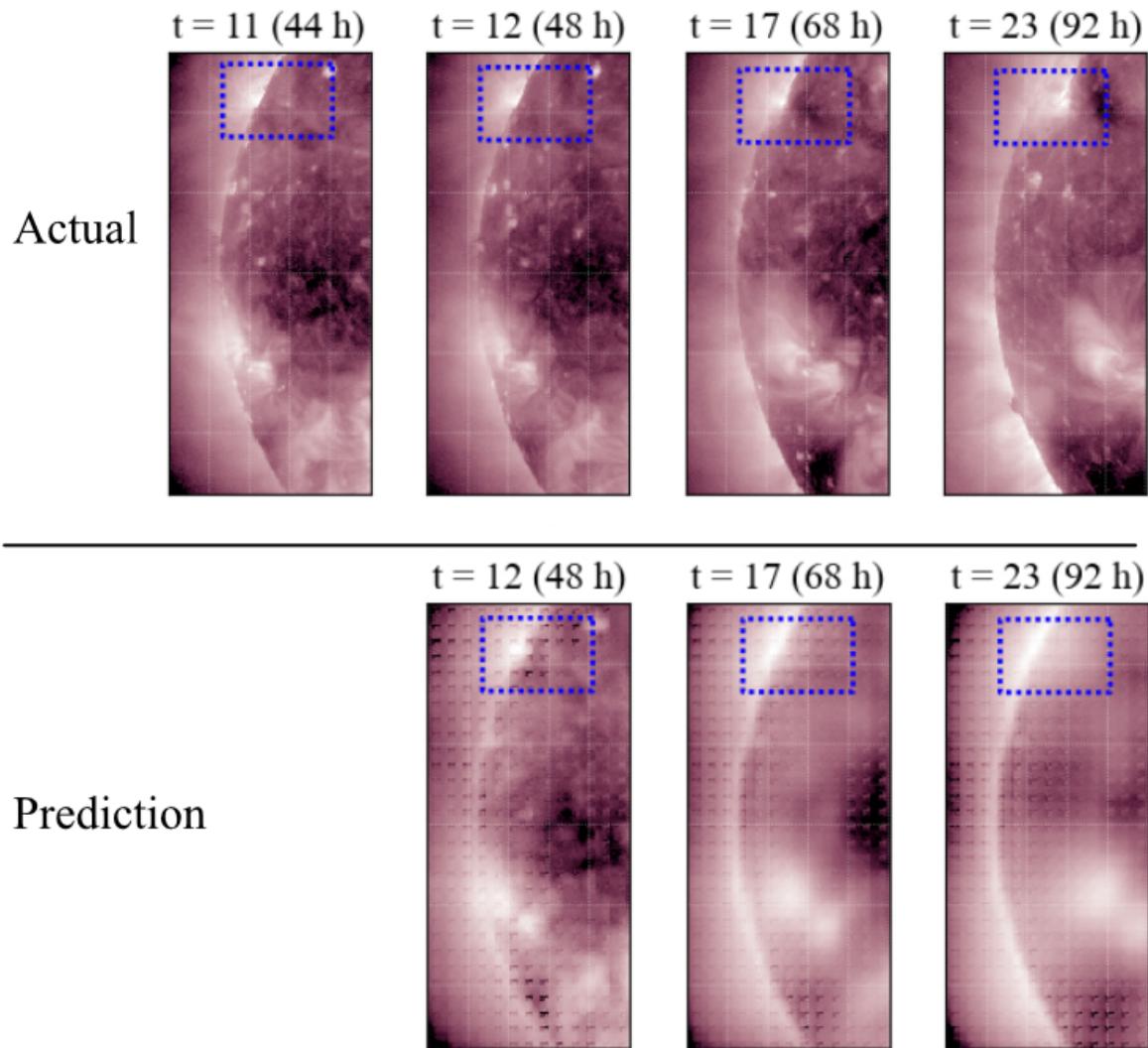


図 5.10: 東側外縁部の北側中緯度帯から出現する活動領域をもつテストセットの例。上段が実際の観測画像、下段がその予測画像である。活動領域を青色破線のバウンディングボックスで囲んでいる。2022年11月6日0時から2022年11月9日8時の期間の画像。

5.4.4 まとめ

5.5 考察

5.5. 考察

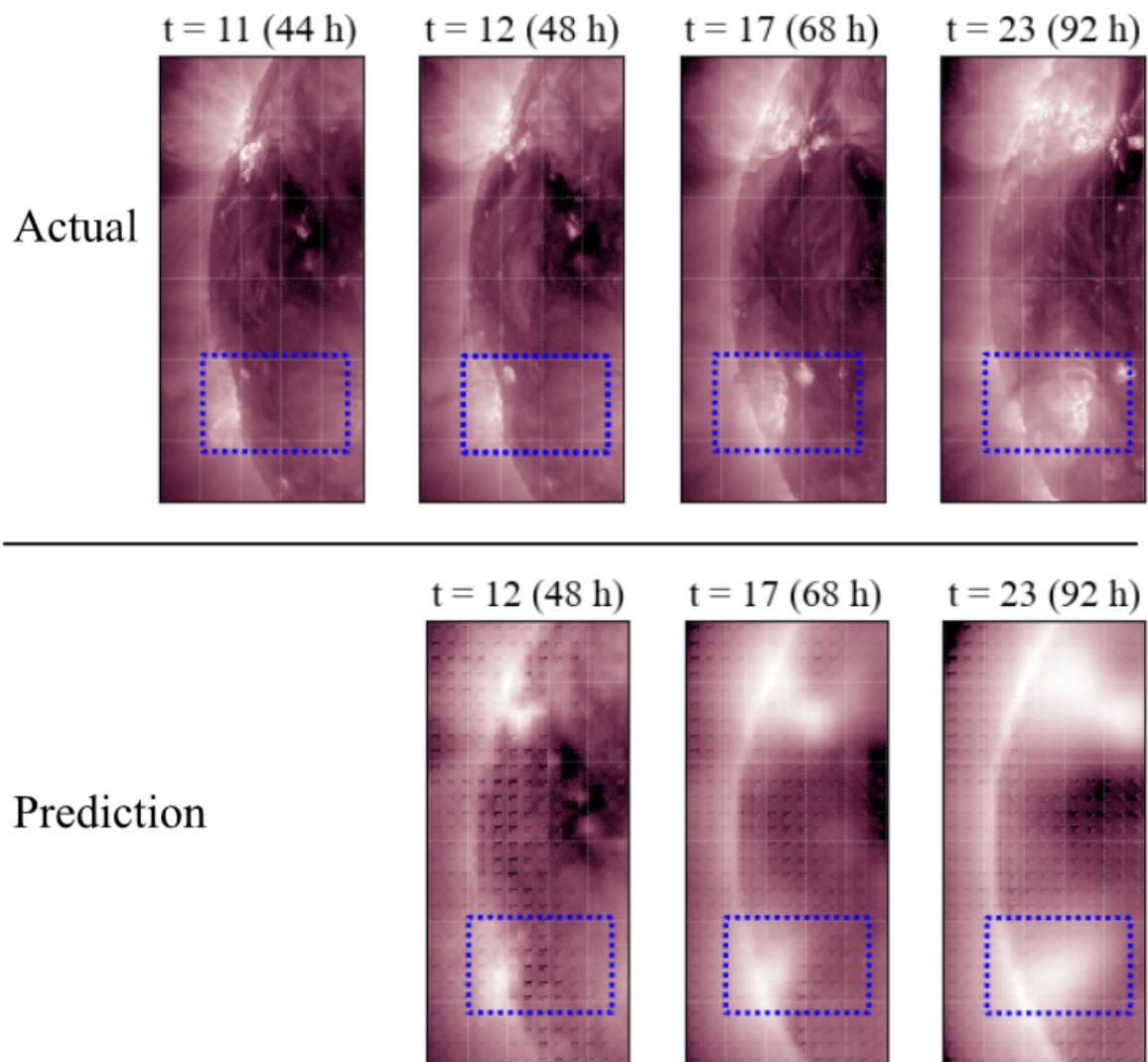
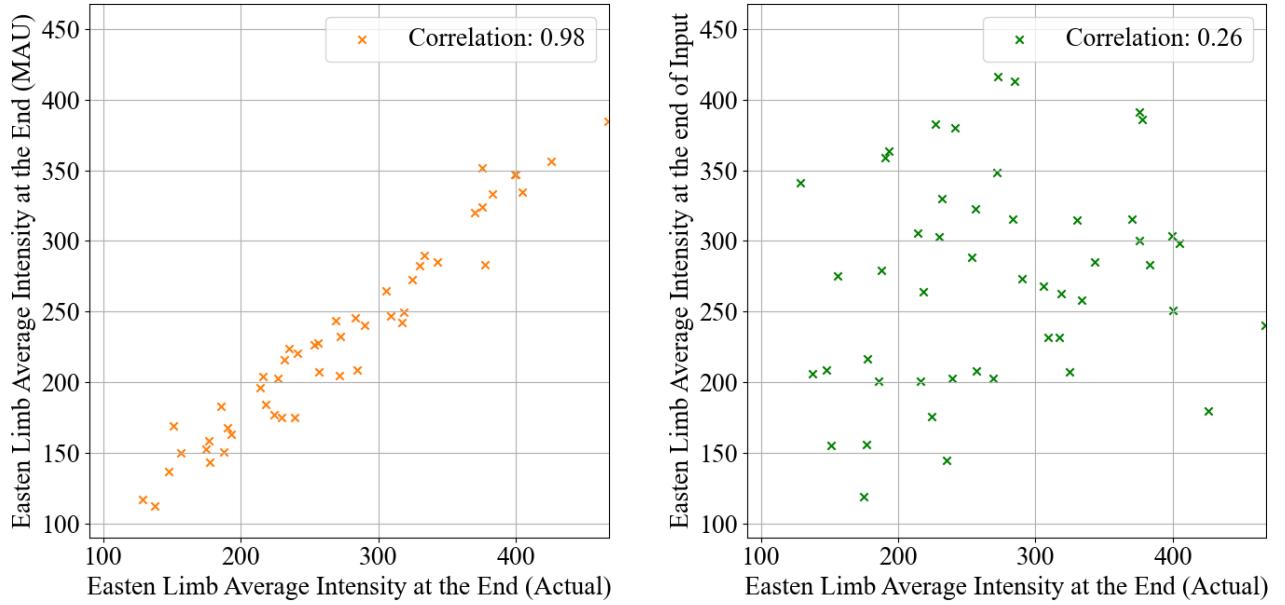


図 5.11: 東側外縁部の南側中緯度帯から出現する活動領域をもつテストセットの例。上段が実際の観測画像、下段がその予測画像である。活動領域を青色破線のバウンディングボックスで囲んでいる。2022 年 12 月 12 日 0 時から 2022 年 12 月 15 日 8 時の期間の画像。

5.5. 考察



(a) すべてのテストセットの、最終タイムステップでの東側 (b) すべてのテストセットでの、最終タイムステップでの東外縁部の平均輝度の予測対実測の散布図。横軸が実際の観側外縁部の平均輝度の実測値と、その48時間前の実測値の測画像から計算された平均輝度強度、縦軸がMAUによる散布図。横軸が実際の観測画像から計算された平均輝度強度予測から計算された平均輝度強度を表す。計算された相関度、縦軸がMAUによる予測から計算された平均輝度強度係数は0.98である。

(b) すべてのテストセットでの、最終タイムステップでの東外縁部の平均輝度の予測対実測の散布図。横軸が実際の観側外縁部の平均輝度の実測値と、その48時間前の実測値の測画像から計算された平均輝度強度、縦軸がMAUによる予測から計算された平均輝度強度を表す。計算された相関度は0.26である。

図 5.12

表 5.2: 本実験での各評価の結果。MAU は、本研究で使用した動画予測モデルによる予測に対する評価、Howard (1990) は、単純差動回転モデルによるシミュレーションに対する評価を表す。

評価指標	全球					経度ごと				
	-90 to -54		-54 to -18		-18 to 18		18 to 54		54 to 90	
平均輝度絶対誤差 ↓										
MAU - 1 波長	0.05	0.04	0.03	0.05	0.06	0.04				
MAU - 3 波長	0.9	0.88	0.89	0.87	0.85	0.86				
Howard (1990)	0.06	0.05	0.04	0.06	0.07	0.05				
SSIM ↑										
MAU - 1 波長	0.9	0.88	0.89	0.87	0.85	0.86				
MAU - 3 波長	0.9	0.88	0.89	0.87	0.85	0.86				
Howard (1990)	0.85	0.87	0.86	0.84	0.83	0.85				

第 6 章

まとめ

謝辞

参考文献

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.
- Zheng Chang, Xinfeng Zhang, Shanshe Wang, Siwei Ma, Yan Ye, Xiang Xinguang, and Wen Gao. Mau: A motion-aware unit for video prediction and beyond. *Advances in Neural Information Processing Systems*, 34:26950–26962, 2021.
- Kunihiro Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202, 1980.
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- RF Howard, JW Harvey, and S Forgach. Solar surface velocity fields determined from small magnetic features. *Solar Physics*, 130:295–311, 1990.
- Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.
- Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
- James R Lemen, Alan M Title, David J Akin, Paul F Boerner, Catherine Chou, Jerry F Drake, Dexter W Duncan, Christopher G Edwards, Frank M Friedlaender, Gary F Heyman, et al. The atmospheric imaging assembly (aia) on the solar dynamics observatory (sdo). *Solar Physics*, 275:17–40, 2012.
- Paulius Micikevicius, Sharan Narang, Jonah Alben, Gregory Diamos, Erich Elsen, David Garcia, Boris Ginsburg, Michael Houston, Oleksii Kuchaiev, Ganesh Venkatesh, et al. Mixed precision training. *arXiv preprint arXiv:1710.03740*, 2017.

参考文献

- Naoto Nishizuka, Komei Sugiura, Yuki Kubo, Mitsue Den, and Mamoru Ishii. Deep flare net (defn) model for solar flare prediction. *The Astrophysical Journal*, 858(2):113, 2018.
- W Dean Pesnell, B J Thompson, and PC Chamberlin. *The solar dynamics observatory (SDO)*. Springer, 2012.
- Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28, 2015.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- Yunbo Wang, Mingsheng Long, Jianmin Wang, Zhifeng Gao, and Philip S Yu. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. *Advances in neural information processing systems*, 30, 2017.
- Yunbo Wang, Lu Jiang, Ming-Hsuan Yang, Li-Jia Li, Mingsheng Long, and Li Fei-Fei. Eidetic 3d lstm: A model for video prediction and beyond. In *International conference on learning representations*, 2018.
- Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- Paul J Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, 1990.