A Deep Physical Model for Solar Irradiance Forecasting with Fisheye Images

Vincent Le Guen [1,2], Nicolas Thome [2]
[1] EDF R&D, Chatou, France
[2] CEDRIC, Conservatoire National des Arts et Métiers, Paris, France

Abstract

We present a new deep learning approach for short-term solar irradiance forecasting based on fisheye images. Our architecture, based on recent works on video prediction with partial differential equations, extracts spatio-temporal features modelling cloud motion to accurately anticipate future solar irradiance. Our method obtains state-of-the-art results on video prediction and 5min-ahead irradiance forecasting against strong recent baselines, highlighting the benefits of incorporating physical knowledge in deep models for real-world physical process forecasting.

1. Introduction

Solar energy is one of the most promising source of renewable energy. Photovoltaic (PV) power generation is steadily increasing worldwide. However, its integration at large scale in the electricity grid is still challenging due to the variable nature of the solar resource, which can lead to stability problems in the connected electricity networks in case of sharp production variations. To overcome the limited spatial and temporal resolution of satellite imaging and numerical weather predictions, sky images from ground-based fisheye cameras have been widely used in recent years [1, 8]. In this paper, we focus on forecasting Global Horizontal Irradiance (GHI), the standard measure of total solar radiation closely related to the output PV power.

Seminal methods for PV forecasting with fisheye images rely on traditional image processing and machine learning techniques. The typical pipeline [4, 2] involves fisheye camera calibration, handcrafted feature extraction, and regression algorithms (nearest neighbors, gradient boosting, multi-layer perceptrons,...) . Cloud motion is often determined with optical flow and extrapolated into the future. Although simple and interpretable, these methods require extensive manual tuning, and poorly adapt to different cameras or site locations. Recently, deep learning models proposed to replace this pipeline with end-to-end data-driven training [16, 11].

Another related line of work is video prediction, for which deep models based on convolutional and/or recurrent architectures [15] have become state-of-the-art. However, natural video forecasting is still a hard high-dimensional extrapolation problem for data-driven algorithms. Generative Adversarial Networks (GANs) [9] were investigated for sharper predictions, and disentangling approaches were proposed (*e.g.* content/motion) [12] to reduce the forecasting dimensionality. Another appealing solution is to inject prior physical knowledge, such as advection-diffusion equation [3], or more general classes of partial differential equations [7, 6].

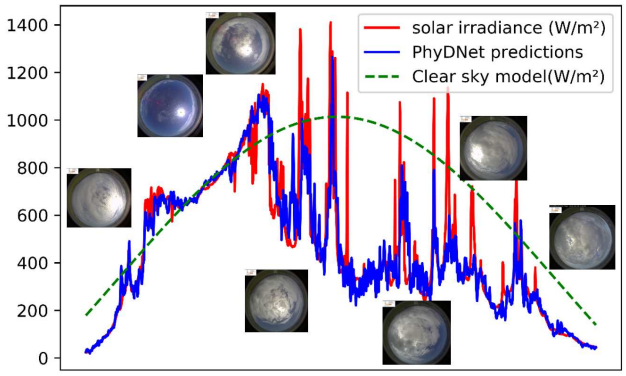

Figure 1. 5min ahead solar irradiance forecasts from fisheye images. Our proposed deep model leveraging physical prior knowledge accurately predicts the sharp intra-day solar irradiance fluctuations.

Predicting future fisheye images is a very challenging task: clouds are deformable objects with complex stochastic behaviour (that can appear or evaporate), several layers with different speeds and directions may be simultaneously present, and the fisheye camera distortion exacerbates the difficulty.

In this work, we introduce a deep neural network model to forecast solar irradiance directly from fisheye images, without any geometric rectification step. Our method, based on PhyDNet [6], exploits physical dynamics to enhance cloud motion modelling. This is to the best of our knowledge the first time physically-constrained models are applied on fisheye images.

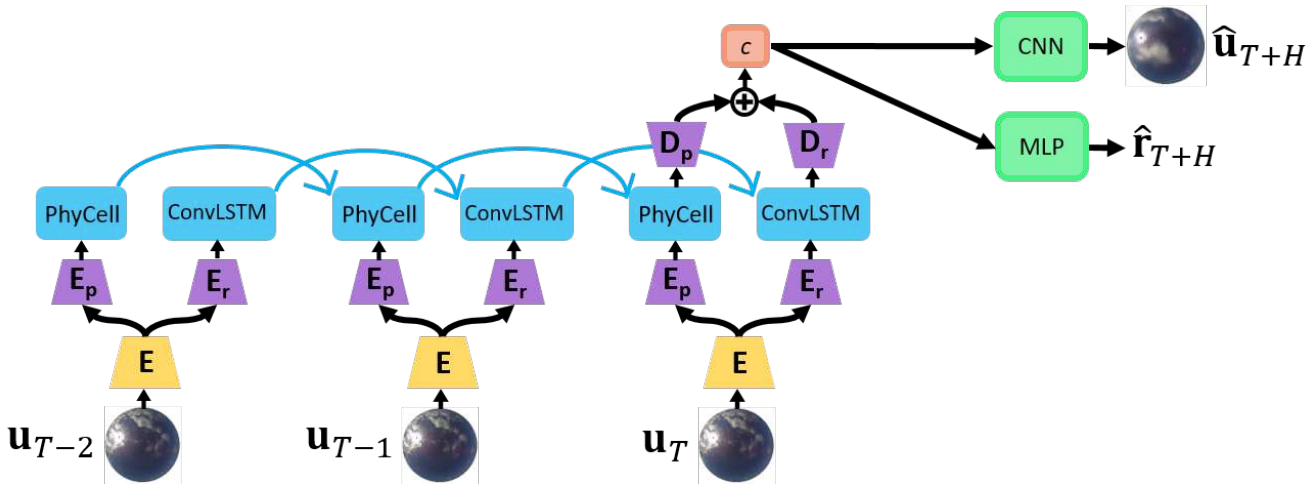

Figure 2. **Proposed deep architecture for solar irradiance forecasting.** Input images are embedded by an encoder $\mathbf{E}$ in a common latent space, followed by specific encoders $\mathbf{E_p}$ and $\mathbf{E_r}$ for extracting physical and residual features. PhyDNet recurrent model is unfolded in time and computes a context vector $c = \mathbf{D_p}(\mathbf{h}_T^{\mathbf{p}}) + \mathbf{D_r}(\mathbf{h}_T^{\mathbf{r}})$, which is used for predicting future irradiance $\hat{r}_{T+H}$ and image $\hat{\mathbf{u}}_{T+H}$.

2. Proposed forecasting model

Given a dataset of fisheye images $\mathbf{u}_{1:T} = (\mathbf{u}_1, ..., \mathbf{u}_T)$ and associated solar irradiance measurements r_t, our goal is to forecast the future irradiance r_{T+H} for a given horizon H. Our architecture builds upon a modified version of the recently proposed PhyDNet model [6] for video prediction.

2.1. Review of PhyDNet model

PhyDNet [6] is a deep architecture that leverages partial differential equations (PDEs) for video prediction. Since physics alone is not sufficient for accurate predictions at the pixel level, PhyDNet aims at learning a latent space $\mathcal{H}$ that linearly disentangles physical dynamics from residual factors (such as texture, details,...). The latent state $\mathbf{h}$ is decomposed into physical and residual components $\mathbf{h} = \mathbf{h^p} + \mathbf{h^r}$, and follows the dynamics:

$$\frac{\partial \mathbf{h}(t,\mathbf{x})}{\partial t} = \frac{\partial \mathbf{h^p}}{\partial t} + \frac{\partial \mathbf{h^r}}{\partial t} := \boldsymbol{\mathcal{M}}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u})) + \boldsymbol{\mathcal{M}}_r(\mathbf{h^r}, \mathbf{E}(\mathbf{u})) \tag{1}$$

The physical model $\boldsymbol{\mathcal{M}}_p$ is composed of a PDE in latent space $\Phi_p(\mathbf{h^p})$ and a correction term $\mathcal{C}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u}))$ with input data (embedded by encoder $\mathbf{E}$): $\boldsymbol{\mathcal{M}}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u})) = \Phi_p(\mathbf{h^p}) + \mathcal{C}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u}))$. The physical predictor Φ_p encodes a general class of linear PDEs up to a differential order q:

$$\Phi_p(\mathbf{h^p}(t,\mathbf{x})) = \sum_{i,j:i+j\leq q} c_{i,j} \frac{\partial^{i+j} \mathbf{h^p}}{\partial x^i \partial y^j}(t,\mathbf{x}) \tag{2}$$

Partial derivatives are computed by constrained convolutions as in PDE-Net [7] and combined by learned coefficients c_{ij}. Discretizing the PDE $\frac{\partial \mathbf{h^p}}{\partial t}(t,\mathbf{x}) = \boldsymbol{\mathcal{M}}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u}))$ with the Euler numerical scheme leads to a recurrent neural network cell (PhyCell). PhyCell performs a physical prediction step in latent space (Eq 3) followed by a correction with embedded input data $\mathbf{E}(\mathbf{u}_t)$ (Eq 4), with a tradeoff controlled by the learned Kalman gain $\mathbf{K}_t$.

$$\tilde{\mathbf{h}}_{t+1}^{\mathbf{p}} = \mathbf{h}_t^{\mathbf{p}} + \Phi_p(\mathbf{h}_t^{\mathbf{p}}) \qquad \textbf{Prediction} \tag{3}$$

$$\mathbf{h}_{t+1}^{\mathbf{p}} = \tilde{\mathbf{h}}_{t+1}^{\mathbf{p}} + \mathbf{K}_t \odot \left(\mathbf{E}(\mathbf{u}_t) - \tilde{\mathbf{h}}_{t+1}^{\mathbf{p}}\right) \qquad \textbf{Correction} \tag{4}$$

The residual model $\boldsymbol{\mathcal{M}}_r(\mathbf{h^p}, \mathbf{E}(\mathbf{u}))$ captures the unknown factors related to unmodelled physics, appearance, texture, and is fully learned from data (implemented by a general ConvLSTM [15]). As shown in [6], any general RNN can also be decomposed as $\boldsymbol{\mathcal{M}}_r(\mathbf{h^r}, \mathbf{u}) = \Phi_r(\mathbf{h^r}) + \mathcal{C}_r(\mathbf{h^r}, \mathbf{E}(\mathbf{u}))$, albeit naively ($\Phi_r(\mathbf{h^r}) = -\mathbf{h^r}(t)$).

2.2. Extension with dual encoders-decoders

One limitation of PhyDNet model is that images $\mathbf{u}_t$ are embedded by an encoder $\mathbf{E}$ in a common latent space for correcting the dynamics of both physical $\mathcal{C}_p(\mathbf{h^p}, \mathbf{E}(\mathbf{u}))$ and residual models $\mathcal{C}_r(\mathbf{h^r}, \mathbf{E}(\mathbf{u}))$. This limits the disentangling ability of PhyDNet since $\mathbf{E}(\mathbf{u}_t)$ contains physical and residual features. We thus propose to learn separate latent spaces for both branches, via additional specific encoders $(\mathbf{E_p}, \mathbf{E_r})$ and decoders $(\mathbf{D_p}, \mathbf{D_r})$, leading to the following model:

$$\frac{\partial \mathbf{h}(t,\mathbf{x})}{\partial t} = \boldsymbol{\mathcal{M}}_p(\mathbf{h^p}, \mathbf{E_p} \circ \mathbf{E}(\mathbf{u})) + \boldsymbol{\mathcal{M}}_r(\mathbf{h^r}, \mathbf{E_r} \circ \mathbf{E}(\mathbf{u})) \tag{5}$$

$\mathbf{E_p}$ aims at learning a specific image embedding for controlling the physical dynamics in latent space with correction features uniquely related to physics (and similarly for $\mathbf{E_r}$). In the following, we denote this model as PhyDNet-dual.

2.3. Architecture for irradiance forecasting

Our proposed architecture, depicted in Figure 2, uses PhyDNet as a physically-constrained RNN for extracting

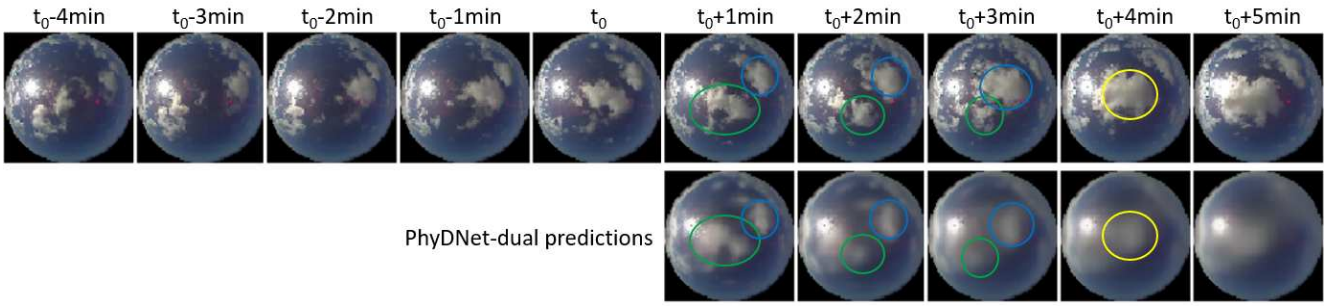

Figure 3. Qualitative fisheye video forecasting results up to 5min horizon. The proposed model successfully predicts the motion of the blue and green clouds that move nearer and finally merge into the yellow cloud.

features from a sequence of past images $\mathbf{u}_{1:T}$. Final physical and residual latent states are first decoded by their respective specific decoders $\mathbf{D_p}$ and $\mathbf{D_r}$ and then summed to get a context vector $c = \mathbf{D_p}(\mathbf{h}^{\mathbf{p}}_T) + \mathbf{D_r}(\mathbf{h}^{\mathbf{r}}_T)$. Then a multi-layer perceptron (MLP) uses the input context c to forecast the future irradiance $\hat{r}_{T+H}$. We also add a CNN to simultaneously forecast the future image $\hat{\mathbf{u}}_{T+H}$. We empirically verified that this multi-task objective improves performances compared to forecasting irradiance only, due to the richer supervision signal and cooperation between tasks (see the first line Table 1).

3. Experimental results

3.1. Fisheye image dataset

A meteorological campaign has been conducted since 2012 at a EDF R&D test site on La Reunion Island. A Fisheye camera (Axis PTZ212) captures whole-sky images every 10s and a pyranometer measures the solar irradiance components GHI and DHI[1]. The whole dataset is composed of more than 6 Million images (resized at resolution 80×80) and associated irradiance measurements. As done classically in the solar energy literature [10], irradiance is normalized with a clear-sky model to remove seasonality and intra-day variation.

3.2. Irradiance forecasting

We forecast solar irradiance at a 5min horizon, given a 5min past context. As illustrated in Figure 1, despite the fast alternation of clouds and sun, we observe that our model successfully anticipates the sharp irradiance fluctuations. We evaluate quantitatively PhyDNet-dual model against recent competitive baselines: ConvLSTM [15] and PredRNN [13], by replacing the video encoder feature extractor by the corresponding RNNs (the context vector c is then the hidden state at the last time step). We report in Table 1 the mean square error (MSE) scores for image prediction $\hat{\mathbf{u}}_{T+5\text{min}}$ and the normalized RMSE[2] for predicted irradiance $\hat{r}_{T+5\text{min}}$.

	image MSE	irradiance nRMSE
PhyDNet irradiance	-	27.8
ConvLSTM [15]	82.7	26.6
PredRNN [13]	82.3	25.1
PhyDNet [6]	80.4	24.4
PhyDNet-dual	**78.8**	**23.5**

Table 1. Solar irradiance and fisheye image 5min forecasting

For both tasks, PhyDNet-dual reaches state-of-the-art performances against all baselines. We also note an improved performance compared to PhyDNet [6], showing that the specific encoders and decoders leads to better spatio-temporal modelling of cloud motion.

3.3. Video prediction

We then evaluate PhyDNet-dual on the video prediction task. Given 5 input images with a 1 min interval, we forecast the 5 future images up to $t_0 + 5$min. We compare PhyDNet-dual with ConvLSTM and Memory In Memory (MIM) [14]. Evaluation metrics are mean squared error (MSE), mean absolute error (MAE) and the structural similarity index SSIM (higher is better). Results shown in Table 2 reveal that PhyDNet-dual outperforms both baselines for all metrics. It confirms that incorporating physical prior information for modelling cloud motion is beneficial compared to fully data-driven algorithms.

	MSE	MAE	SSIM
ConvLSTM [15]	83.1	681	0.845
MIM [14]	68.6	635	0.840
PhyDNet-dual	**68.1**	**629**	**0.862**

Table 2. Quantitative video prediction results.

We show in Figure 3 a video prediction example of

[1] Global and Diffuse Horizontal Irradiance respectively

[2] nRMSE = Root Mean Square Error normalized by the mean value of the quantity on the train set, expressed as a percentage.

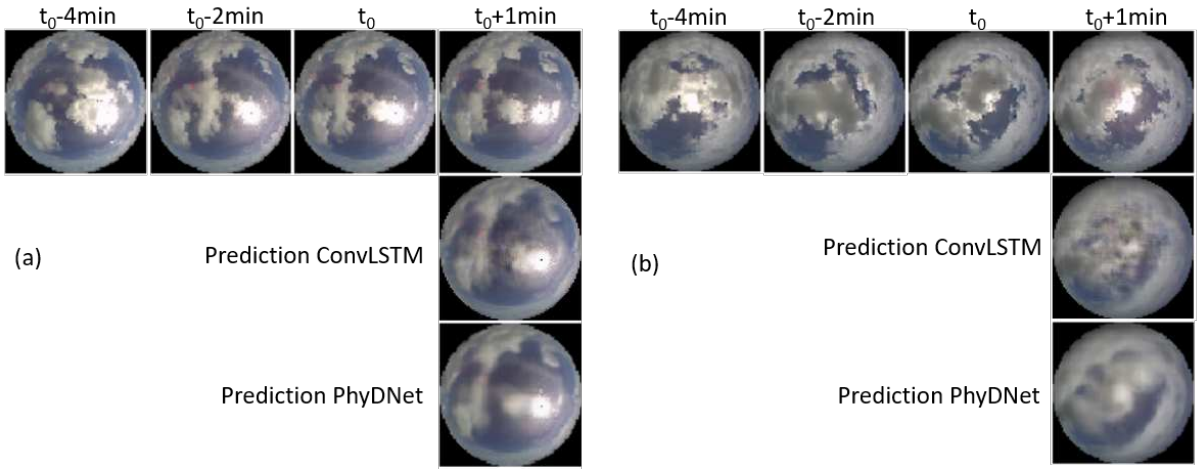

Figure 4. Qualitative forecasting comparison between PhyDNet-dual and ConvLSTM.

PhyDNet-dual model. The future of this sequence presents 2 clouds (circled in blue and green) moving closer between t_0 and $t_0 + 3$min and finally merging at time $t_0 + 4$min. We observe that PhyDNet-dual predicts the same outcome with a good accucary on cloud location, although clouds become blurry because of uncertainty.

In Figure 4, we provide a particular comparison to ConvLSTM [15], which forms the residual branch of PhyDNet. In sequence (a), we see that the shape of the small cloud getting nearer the sun is much better predicted by PhyDNet-dual. In sequence (b), the sun will reappear 1 min in the future. PhyDNet-dual provides a better anticipation by prediction a bright spot at the sun location and better defined cloud shapes. It confirms that incorporating physical dynamics in a deep model brings great improvement for predicting natural phenomena, with a very small amount of additional parameters with respect to ConvLSTM.

4. Conclusion

We tackle solar irradiance forecasting directly from raw omnidirectional fisheye images with a physically-constrained deep model. Our model outperforms recent fully data-driven baselines on a large real-world dataset. Future work include using more specific physical models [3], adequate loss functions [5], and probabilistic forecasting.

References

[1] C. W. Chow, B. Urquhart, M. Lave, and A. e. a. Dominguez. Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed. *Solar Energy*, 2011. 1

[2] Y. Chu, M. Li, and C. Coimbra. Suntracking imaging system for intra-hour DNI forecasts. *Renewable Energy*, 2016. 1

[3] E. de Bezenac, A. Pajot, and P. Gallinari. Deep learning for physical processes: Incorporating prior scientific knowledge. *ICLR*, 2018. 1, 4

[4] C. Gauchet, P. Blanc, B. Espinar, B. Charbonnier, and D. Demengel. Surface solar irradiance estimation with low-cost fish-eye camera. In *COST WIRE*, 2012. 1

[5] V. Le Guen and N. Thome. Shape and time distortion loss for training deep time series forecasting models. *NIPS*, 2019. 4

[6] V. Le Guen and N. Thome. Disentangling physical dynamics from unknown factors for unsupervised video prediction. In *CVPR*. 2020. 1, 2, 3

[7] Z. Long, Y. Lu, X. Ma, and B. Dong. PDE-Net: Learning PDEs from data. In *ICML*, 2018. 1, 2

[8] R. Marquez and C. F. Coimbra. Intra-hour dni forecasting based on cloud tracking image analysis. *Solar Energy*, 91:327–336, 2013. 1

[9] M. Mathieu, C. Couprie, and Y. LeCun. Deep multi-scale video prediction beyond mean square error. In *International Conference on Learning Representations (ICLR)*, 2015. 1

[10] C. Rigollier, O. Bauer, and L. Wald. On the clear sky model of the ESRA (European Solar Radiation Atlas) with respect to the heliosat method. *Solar energy*, 2000. 3

[11] Y. Sun, V. Venugopal, and A. R. Brandt. Short-term solar power forecast with deep learning: Exploring optimal input and output configuration. *Solar Energy*, 2019. 1

[12] R. Villegas, J. Yang, S. Hong, X. Lin, and H. Lee. Decomposing motion and content for natural video sequence prediction. *ICLR*, 2017. 1

[13] Y. Wang, M. Long, J. Wang, Z. Gao, and S. Y. Philip. PredRNN: Recurrent neural networks for predictive learning using spatiotemporal lstms. In *NeurIPS*, 2017. 3

[14] Y. Wang and J. e. a. Zhang. Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. In *CVPR*, 2019. 3

[15] S. Xingjian, Z. Chen, and H. e. a. Wang. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *NIPS*, 2015. 1, 2, 3, 4

[16] J. Zhang, R. Verschae, S. Nobuhara, and J.-F. Lalonde. Deep photovoltaic nowcasting. *Solar Energy*, 2018. 1