Lecture 1b - Introduction

Dr. Yu Zhang

ECE Dept, UC Santa Cruz

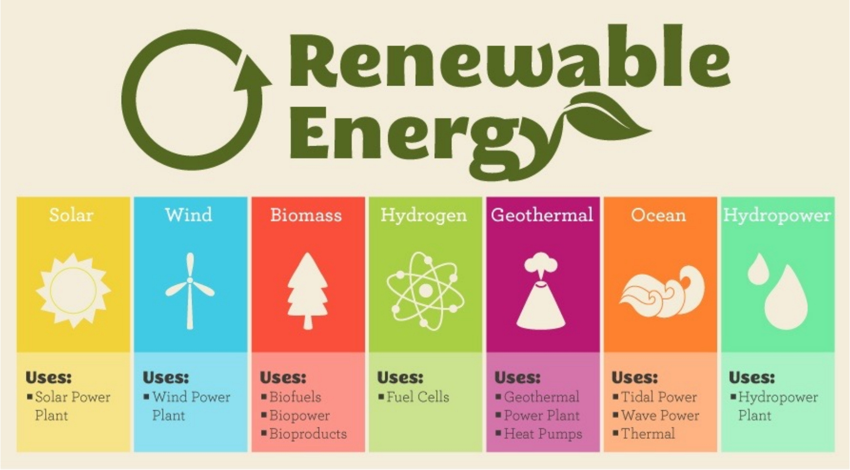
ECE180J

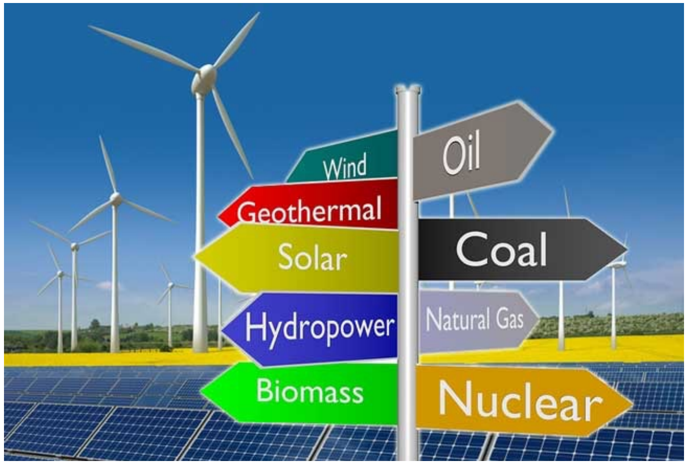
Outline

- Renewables & Non-renewables
- Greenhouse Effect
- Energy Storage
- Smart Power Grids

Renewables vis-à-vis Non-renewables

- **Renewables:** Collected from renewable resources, which are naturally replenished on a human timescale, such as sunlight, wind, rain, tides, waves, and geothermal heat. [Wiki]
- Human timescale: "Anything in space or geological is not on our time scale. Anything to do with our society and living organisms, excluding evolution, is." [Quora]
- **Nonrenewables (finite resource):** Resource that does not renew itself at a sufficient rate for sustainable economic extraction in meaningful human time-frames. e.g. earth minerals and metal ores, fossil fuels (coal, petroleum, natural gas) and groundwater. [Wiki]





Solar Power

Solar power: conversion of energy from sunlight into electricity, either directly using photovoltaics (PV), indirectly using concentrated solar power (CSP), or a combination. [Wiki]

- PV cells convert light into an electric current via the PV effect

- Use lenses/mirrors and tracking systems to focus a large area of sunlight into a small beam.

Panda Solar Panels

248-acre panda solar farm in Datong, Shanxi, China.

• Built by China Merchants New Energy Group, one of the country's largest clean energy operators.

• The 1st phase, which includes one 50 MW plant, was completed on 6/30/17. A second panda is planned.

• It will produce 3.2 billion kWh of solar energy, and reduce carbon emissions by 2.74 million tons in 25 years.

Concentrating Solar Power (CSP)

The Crescent Dunes Solar Energy Project:

- Scale: 110 MW net solar thermal power + 1.1 GWh of energy storage
- Location: near Tonopah, 190 miles northwest of Las Vegas.
- Features: first utility-scale CSP plant with a central receiver tower + advanced molten salt storage tech from SolarReserve.
- Cost less than $1 billion.
- Planned energy output was 500 GWh.

Wind Power

Wind power: Use of air flow through wind turbines to mechanically power generators for electricity. [Wiki]

- Offshore turbines are located out at sea or in freshwater.
- Onshore wind refers to turbines located on land

Wind Turbine

Blade transportation

Turbine explode

Other Renewables

Hydropower

Geothermal energy

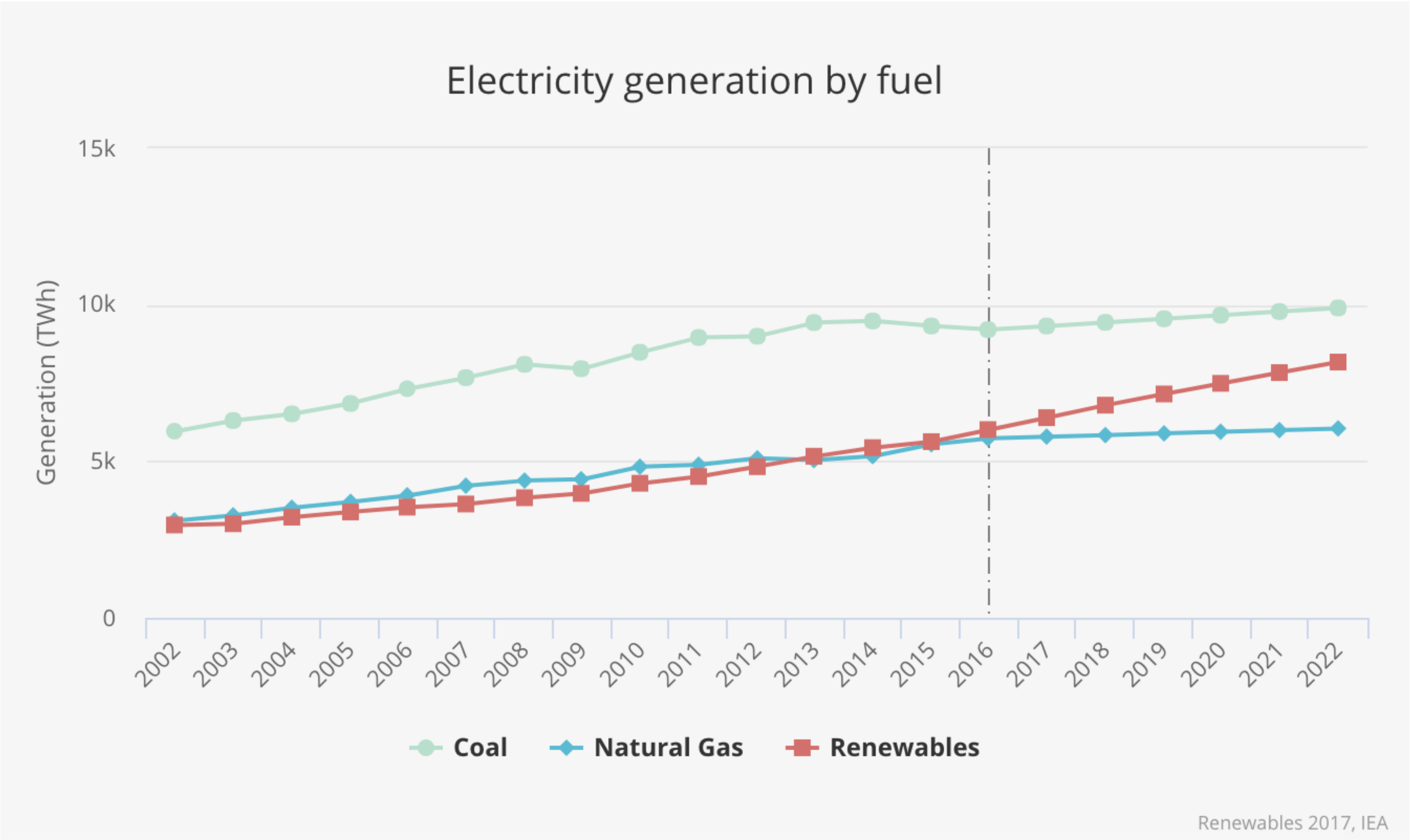
Bioenergy

Tidal energy

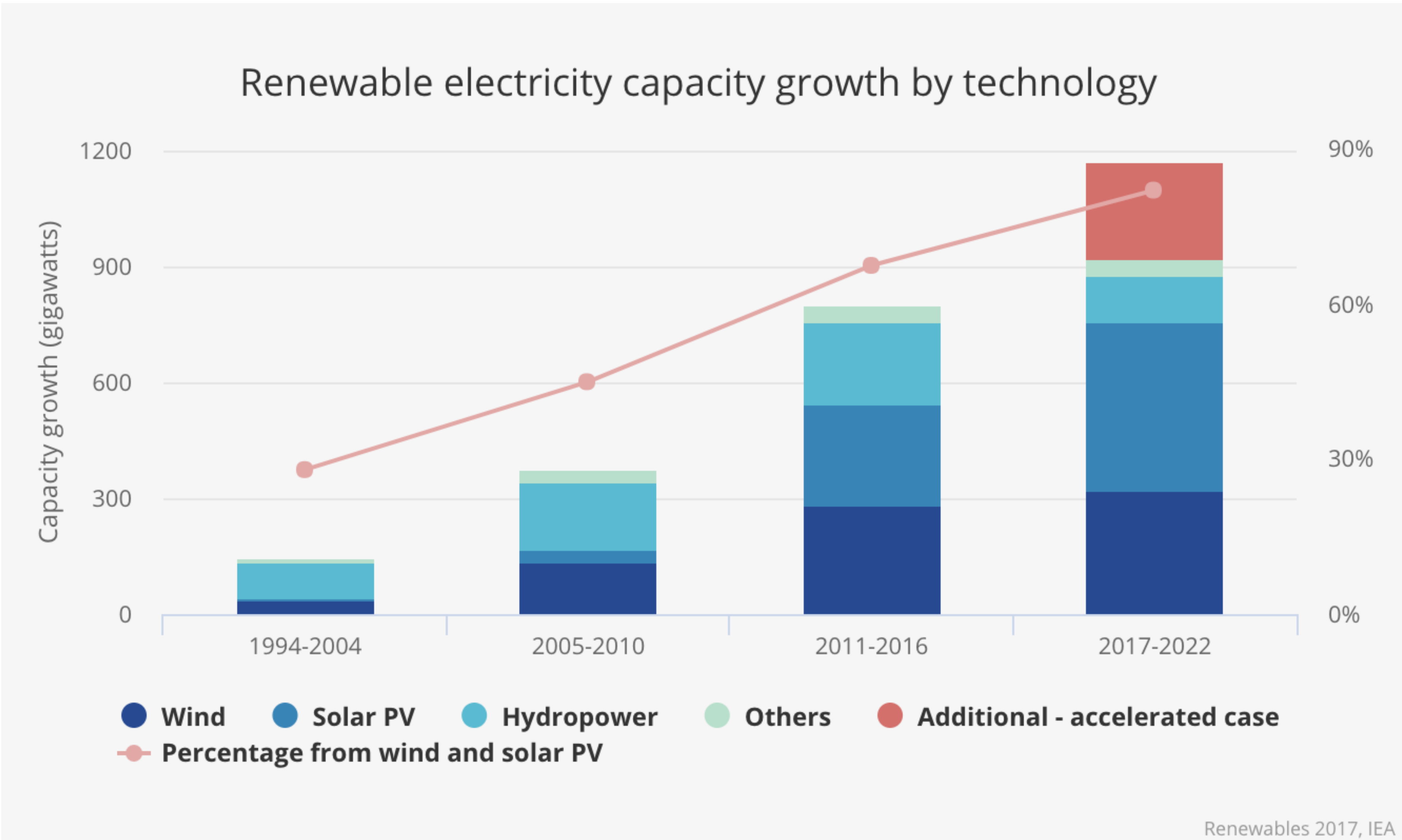
Investments in Renewables

Figure source: www.irena.org

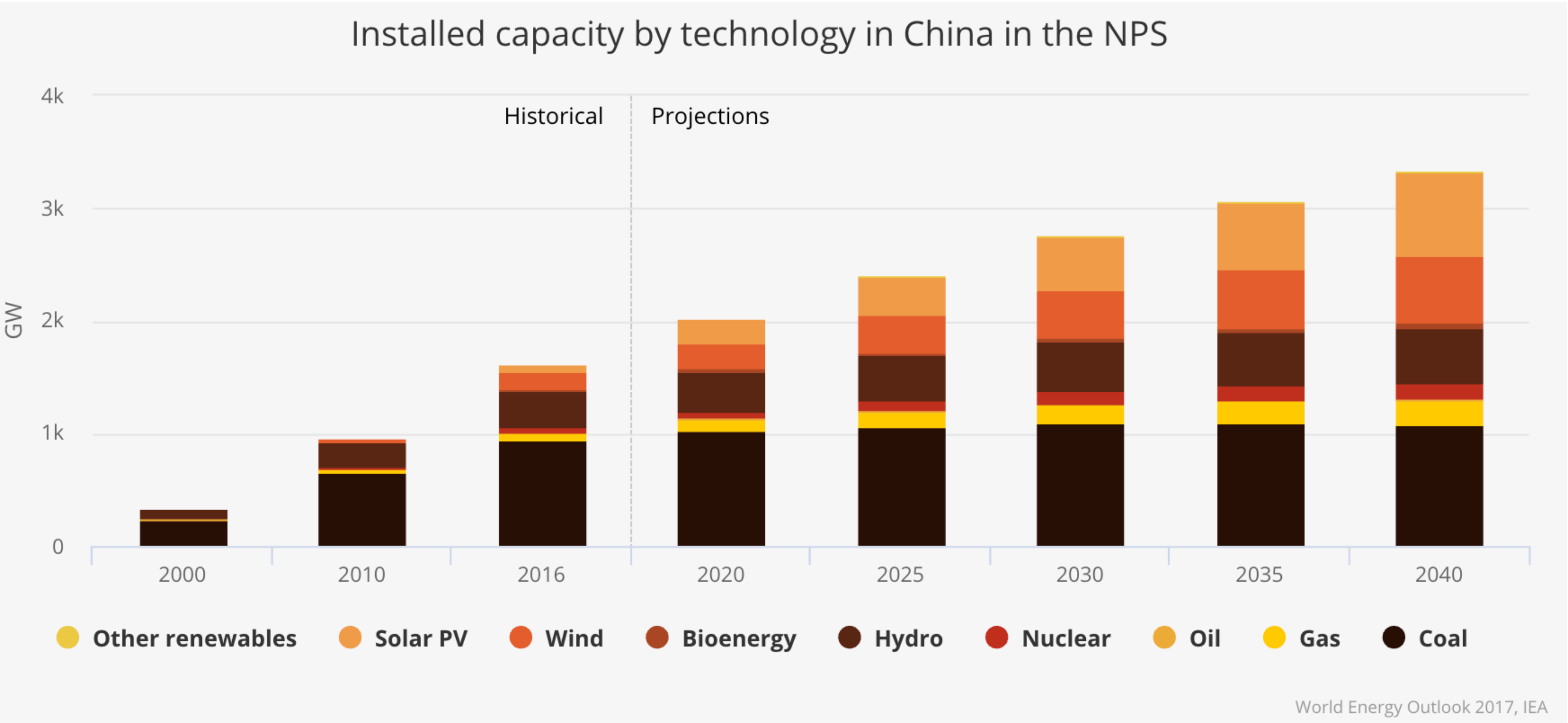
Electricity Generation



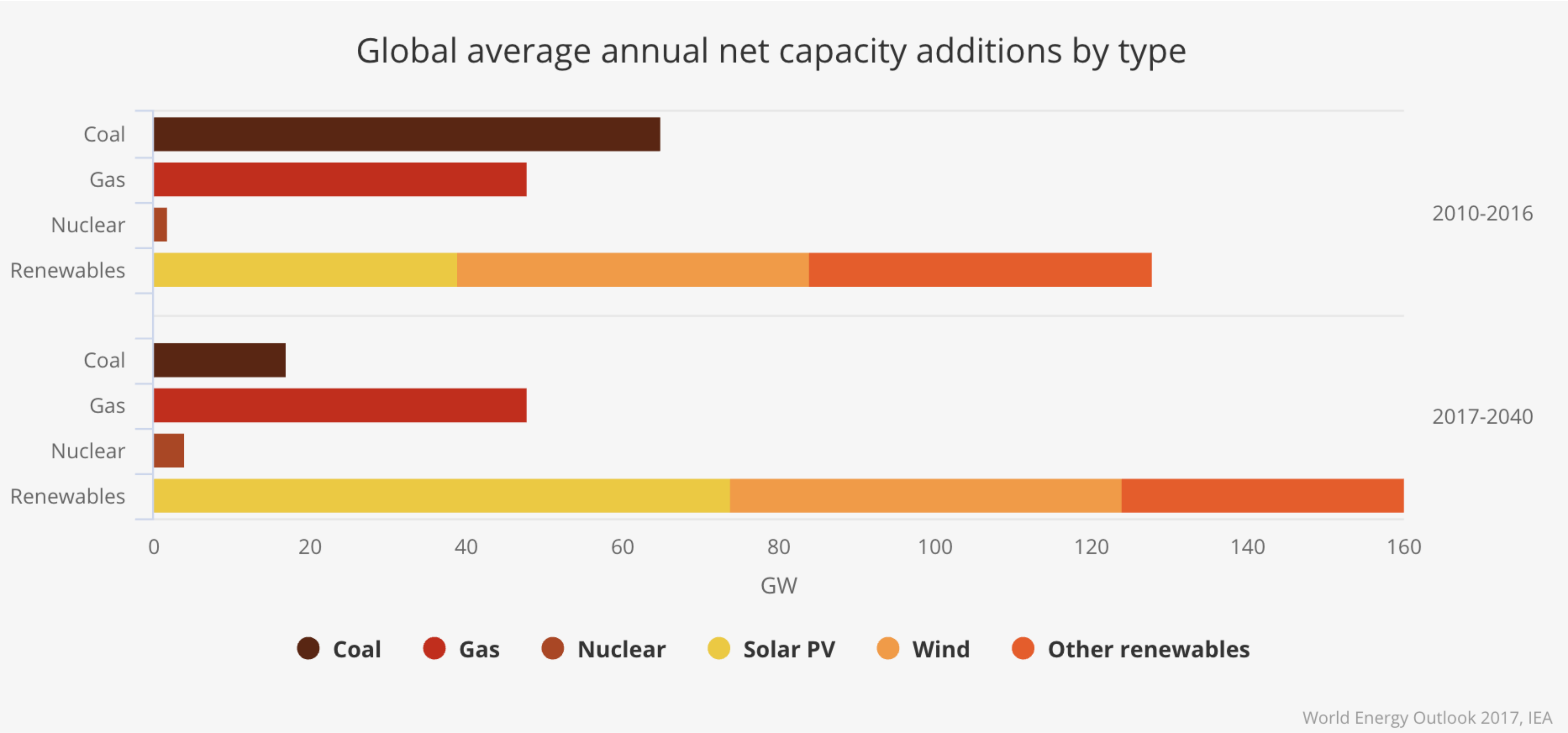
Capacity Growth of Renewables



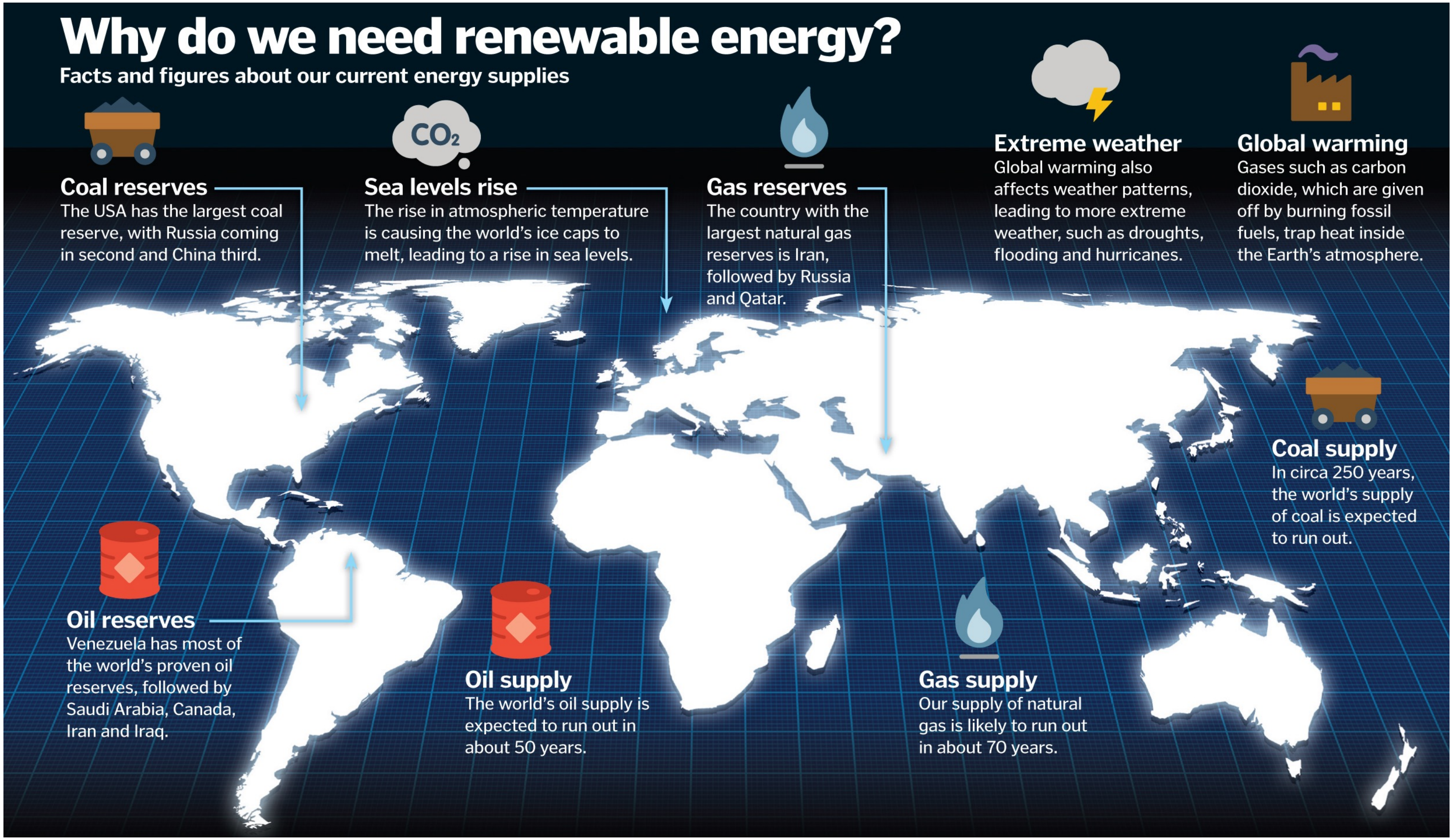
Bright Future of Renewables



Bright Future of Renewables

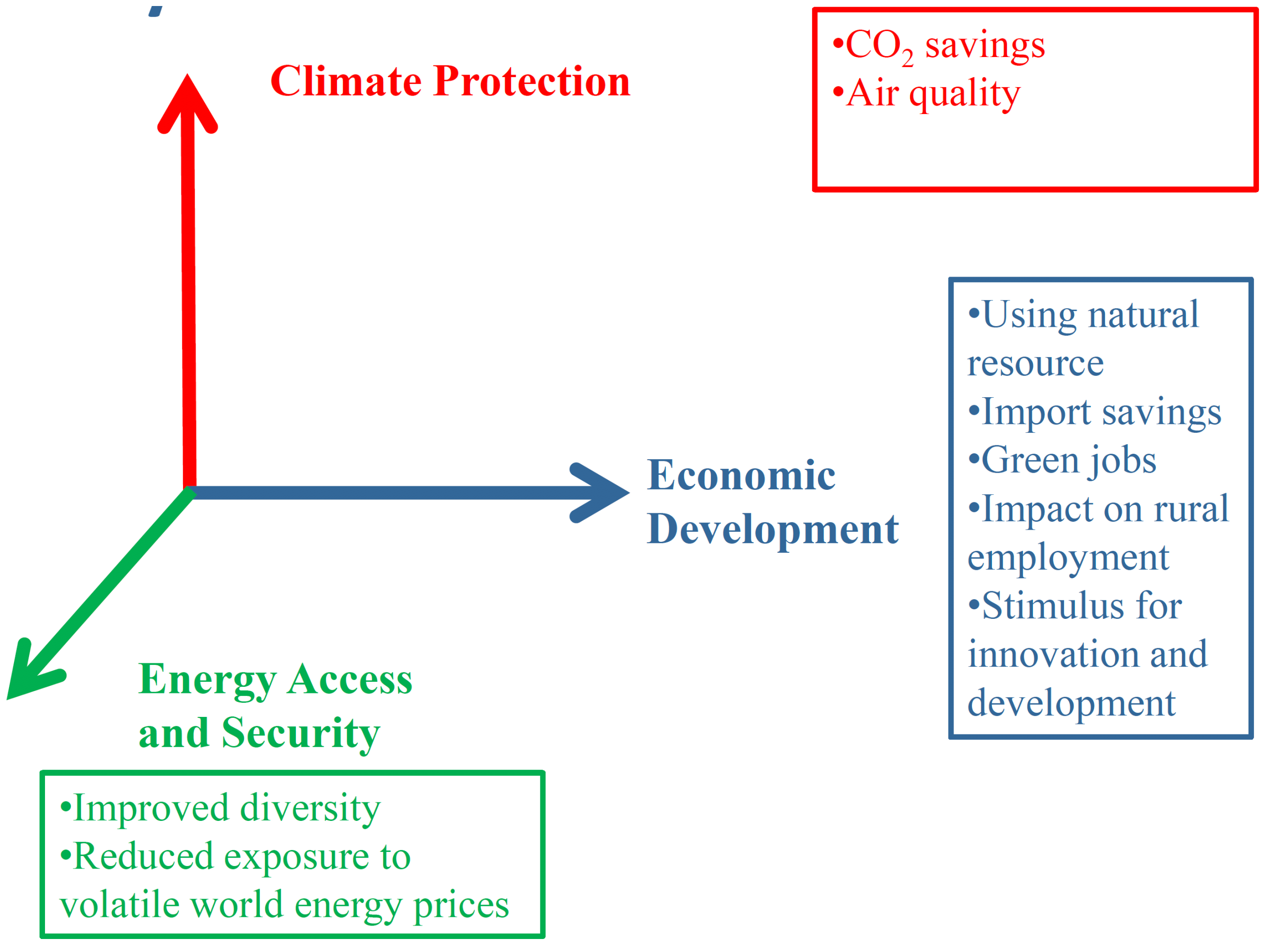


Why Renewables?



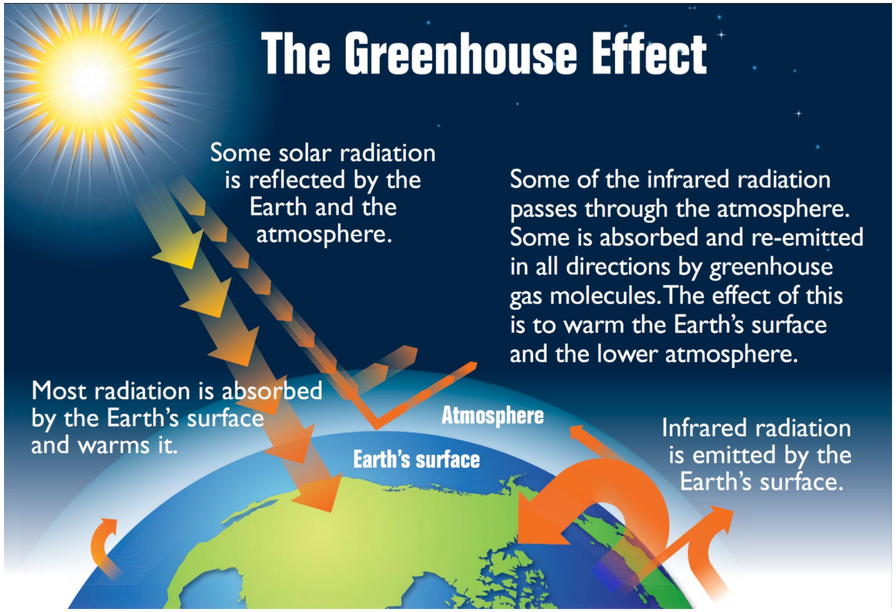
Pic credit: www.howitworksdaily.com/why-do-we-need-renewable-energy

Benefits of Renewables



Pic credit: Richard Jones, Deputy Executive Director Intl Energy Agency

The Greenhouse Effect



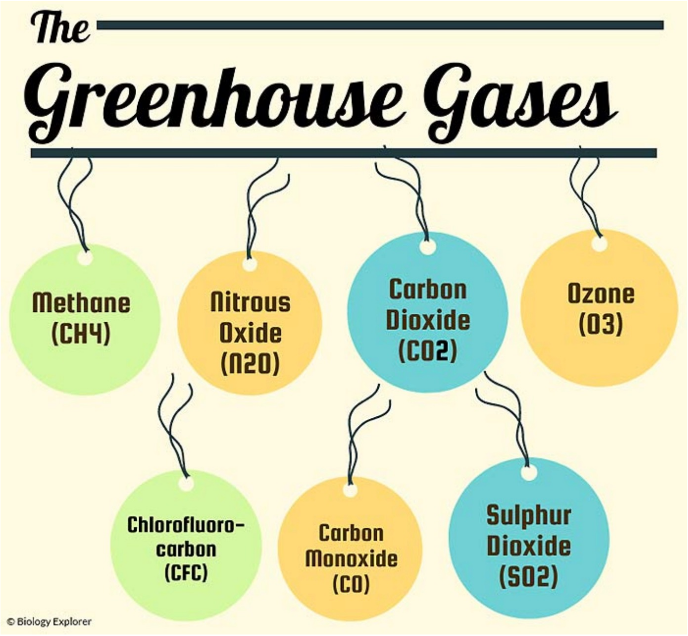
- The sun naturally bombs the Earth with significant amounts of radiation: visible light, infrared (IR) light, ultraviolet (UV) light, etc.
- When sunlight reaches the Earth, some radiation from the sun is absorbed by the surface of the land, whereas some portions are reflected.

- A greenhouse gas is an atmospheric gaseous substance that can sufficiently absorb IR radiation.
- The molecules of these greenhouse gases absorb the heat and reflect them to the surface. Because of this, the temperature in the Earth's atmosphere and land surface increases.
- Naturally, a greenhouse effect is considered an essential phenomenon that helps maintain the Earth's temperature. Without it, the planet's temperature would fall to just -18° C rather than +15° C.

Pic credit: Climate Change Indicators in the USA (2nd ed.), US EPA, Dec 2012.

The Greenhouse Gases (GHGs)

GHGs	Main Source
H_2O	Evaporation of oceans & lakes
CO_2	Combustion of fossil fuels
CH_4	Anaerobic decay of organic matter by livestock
N_2O	Artificial fertilizers
CFCs	Refrigerants and solvents



The increase in the amount of CO_2 accounts for more than half of the greenhouse effect, compared with other gases like CH_4, N_2O, O_3, and CFCs which account for 16%, 5%, 12%, and 12%, respectively.

Effects of Too Much GHGs

1. Global Warming

Changes in the Earth's temperature will trigger changes in wind patterns and variations of cloud covers.

2. Rise in Sea Level

Studies predict that the Earth's sea level will become ten times higher in the next 90 years if global warming continues.

3. Effects on Biological Cycles

Global warming also contributes to the loss of coastal areas and other habitats, thereby reducing organism population and diversity.

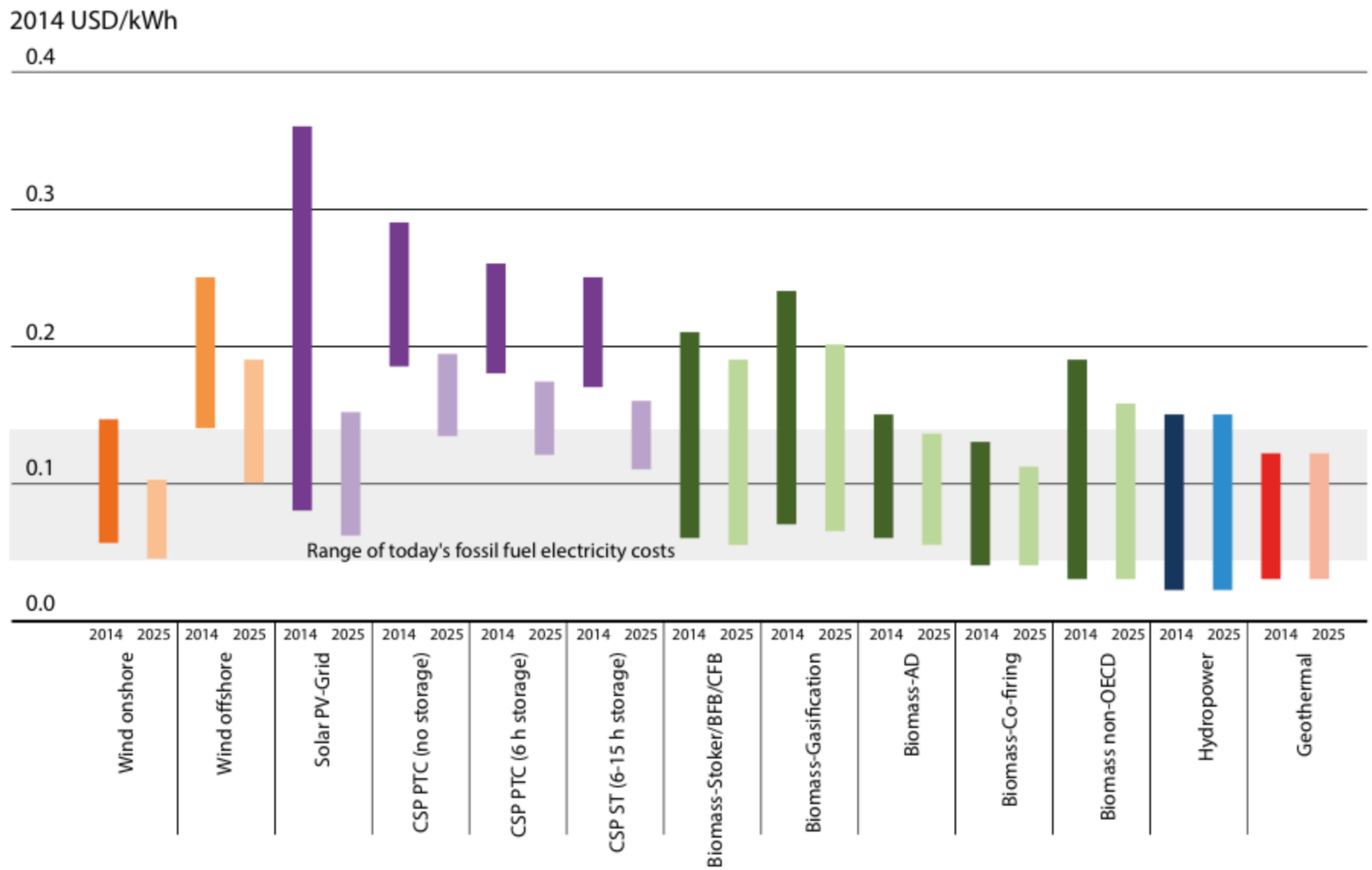
4. Economic and Agricultural Impacts

Too much heat can affect soil and water conditions, causing them to wilt and die.

Source: www.bioexplorer.net/greenhouse-gases.html/

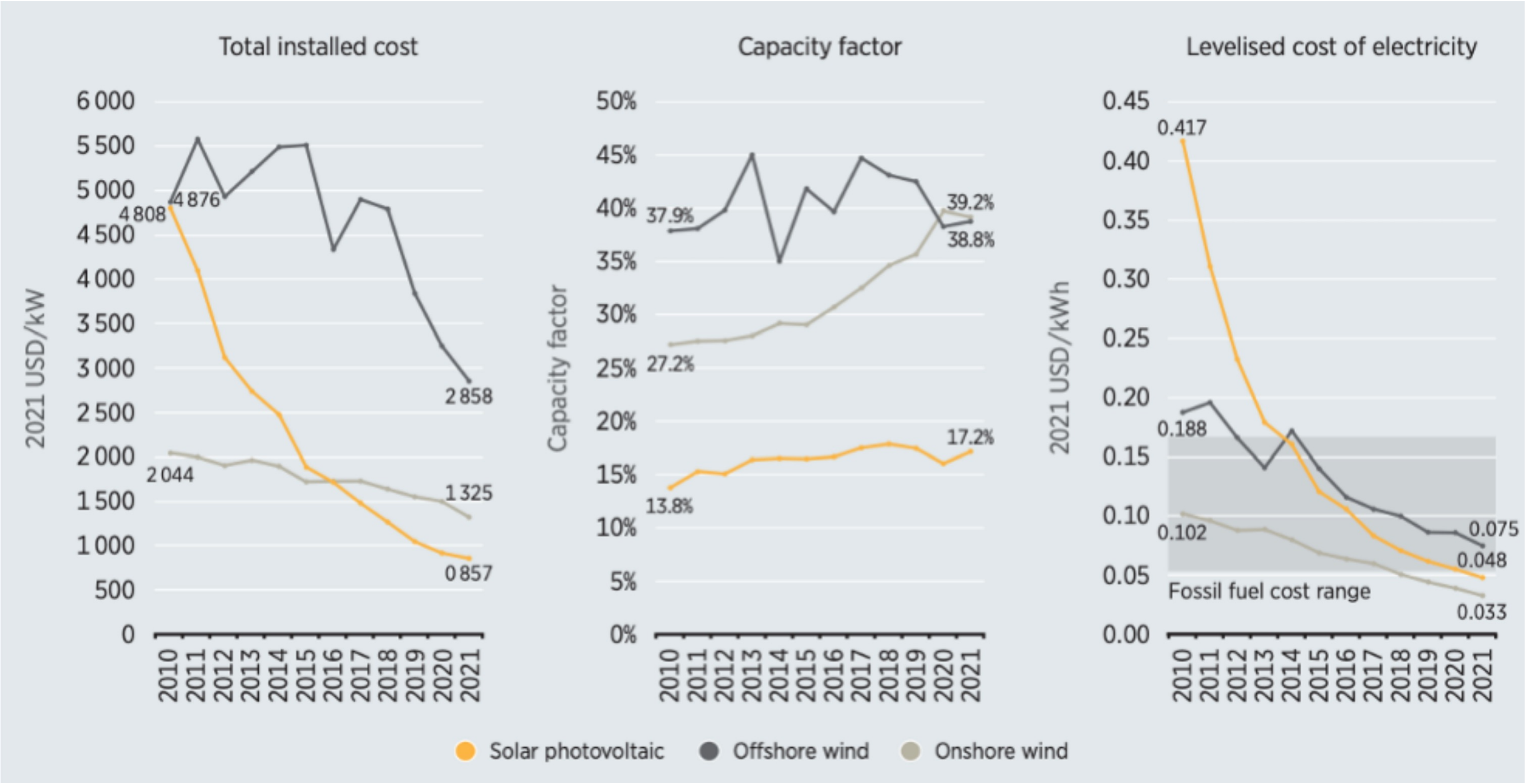
Costs of Renewable Energy

Source: IRENA Renewable Cost Database and Auctions Database.



- PTC (Parabolic Trough Collector) vs ST (Solar Tower)
- AD (Anaerobic Digestion)
- OECD (Organization for Economic Co-operation and Development)

Costs of Renewable Energy

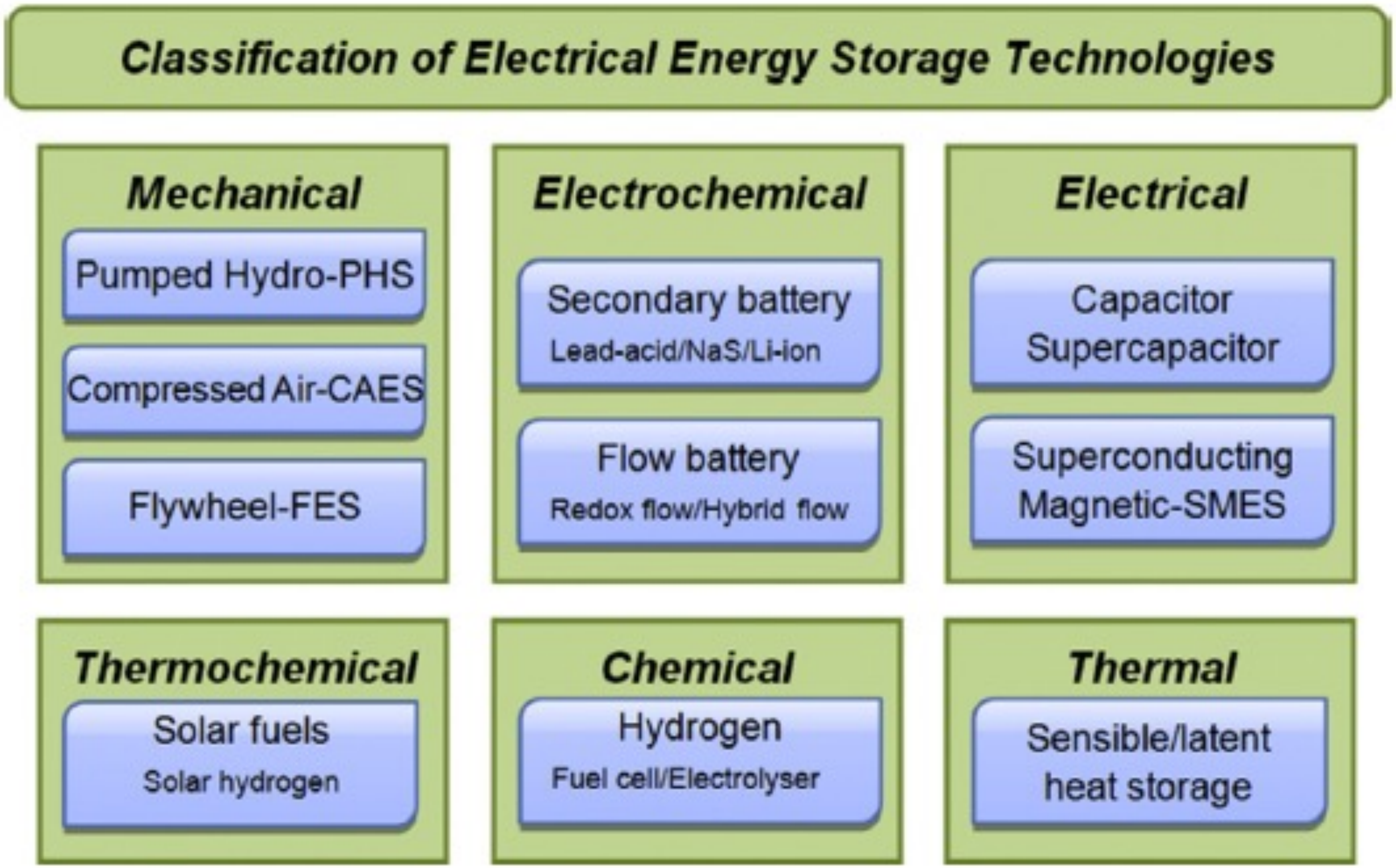


- Capacity factor: Ratio of actual electrical energy output to the theoretical maximum electrical energy output.
- Levelized cost of energy (LCOE): Average net present cost over lifetime of a plant.

Source: IRENA

Energy Storage

Capture of energy produced at one time for use at a later time. A device that stores energy is called an accumulator or battery. [Wiki]

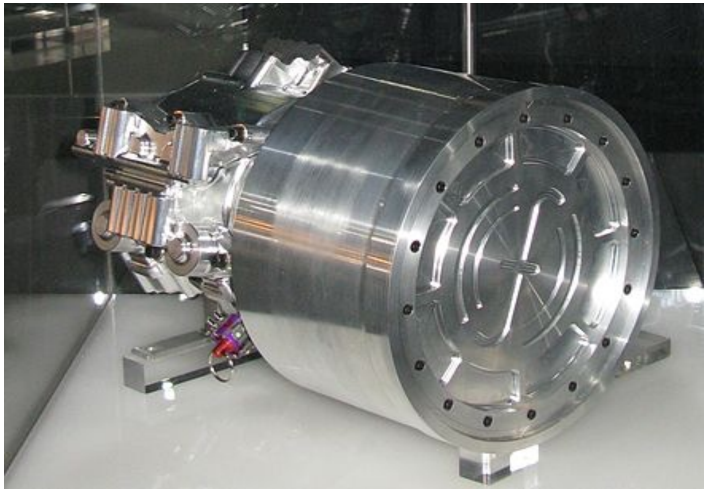


Picture credit: Xing Luo, etc, “Overview of current development in electrical energy storage technologies and the application potential in power system operation.”

Energy Storage

Pumped storage hydroelectricity

Compressed air locomotive

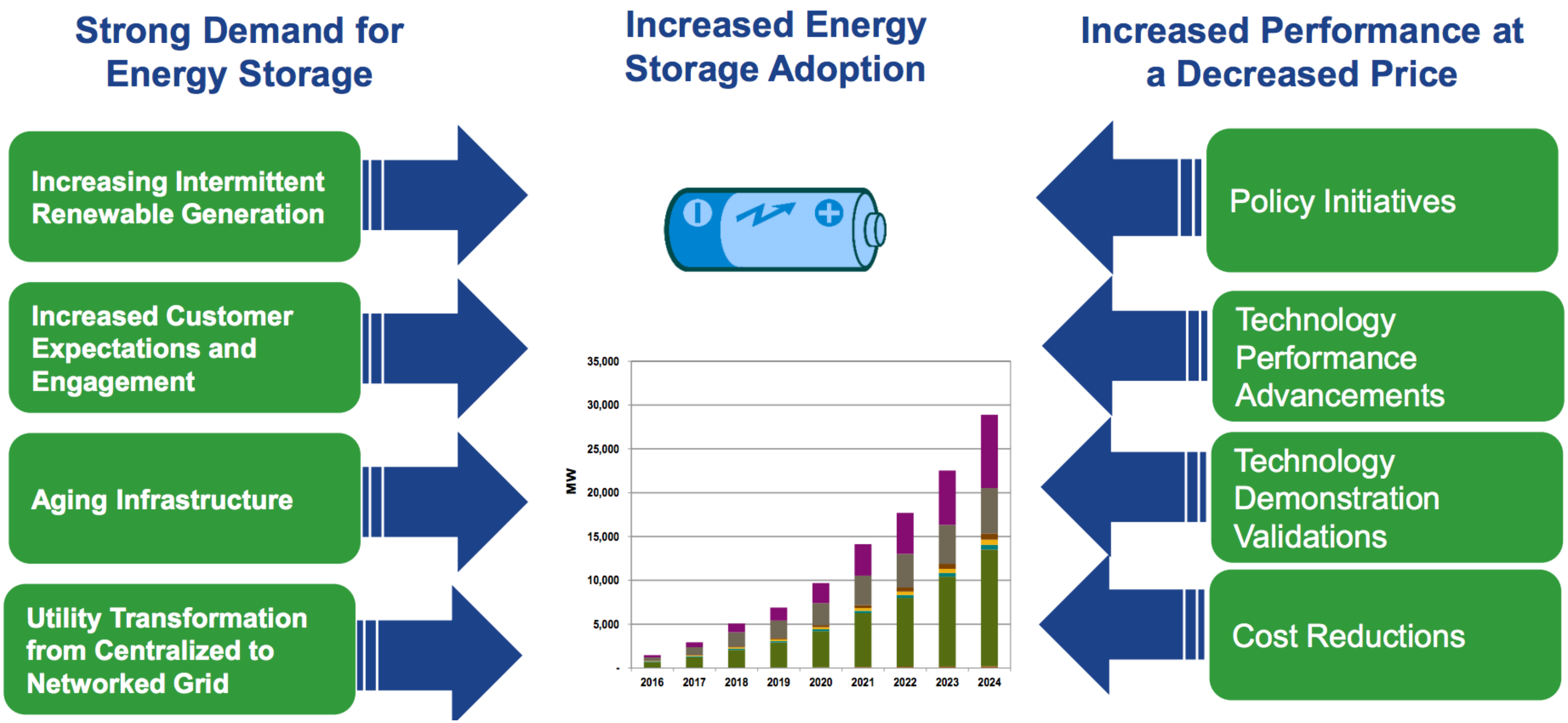


Flywheel

Thermal storage

Why Energy Storage Now

Driving demand: Industry changes, policy, technology, and cost advances



Picture credit: betterbuildingssolutioncenter.energy.gov

Applications of Energy Storage

Electricity cost optimization

- peak/off-peak price management
- demand and power factor charge management

Capacity

- generation resource adequacy (e.g., capacity markets, operating reserves)
- Transmission & distribution infrastructure adequacy

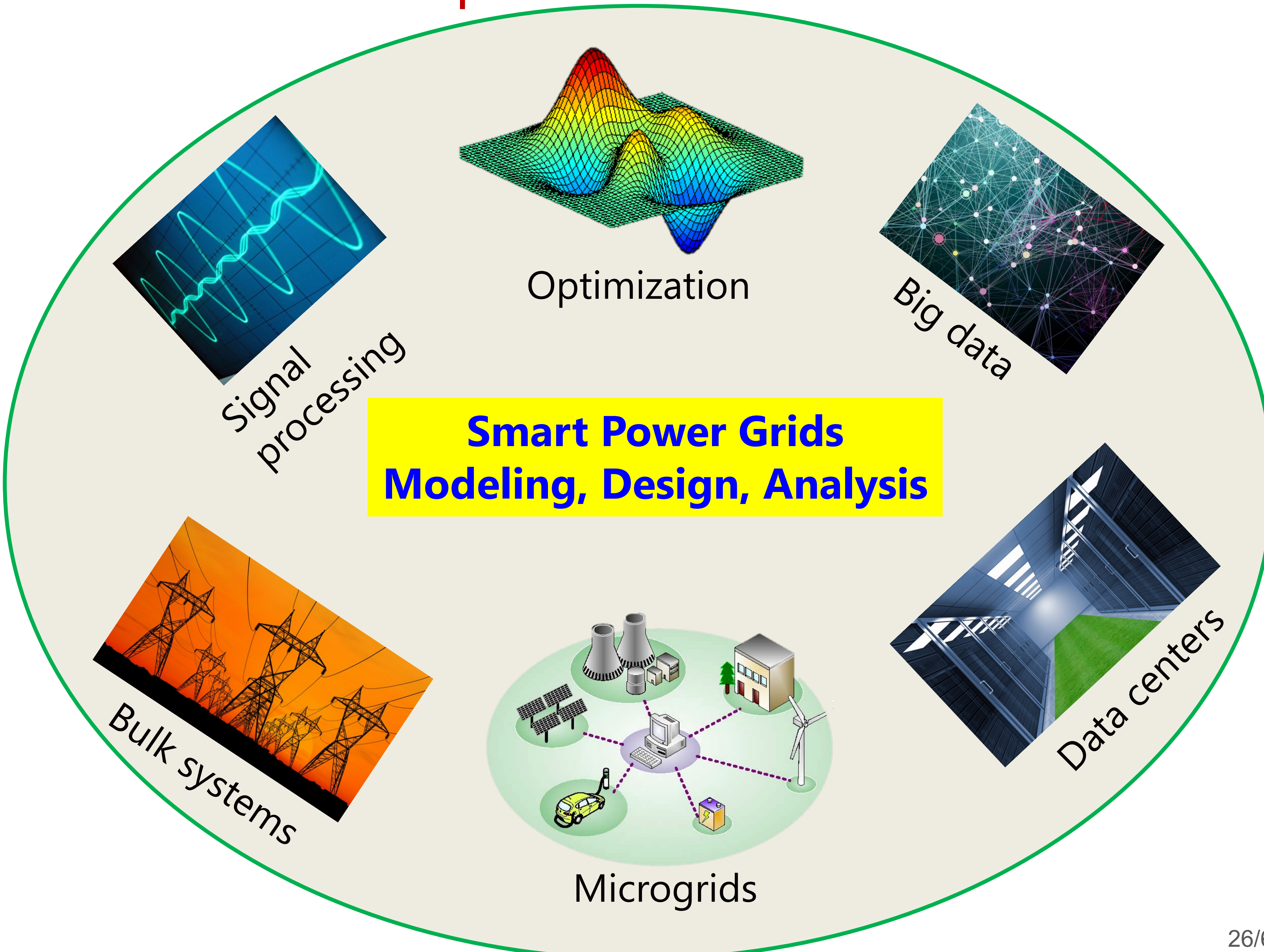
Routine grid operations

- frequency regulation
- voltage/VAR support
- renewable energy ramping/ smoothing/shifting

Contingency Situations

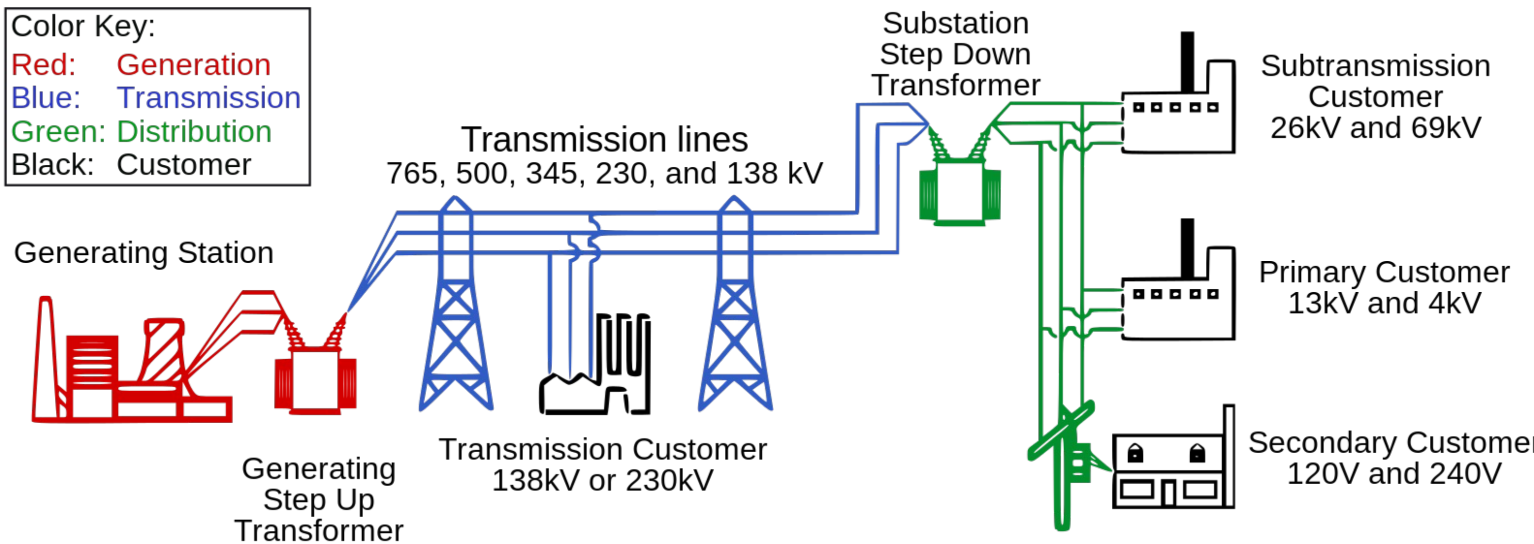
- black start
- sustained/momentary outages

Research Landscape



The Electrical Grid Then and Now

"Most significant engineering achievement of 20th century" [NAE Report'10]

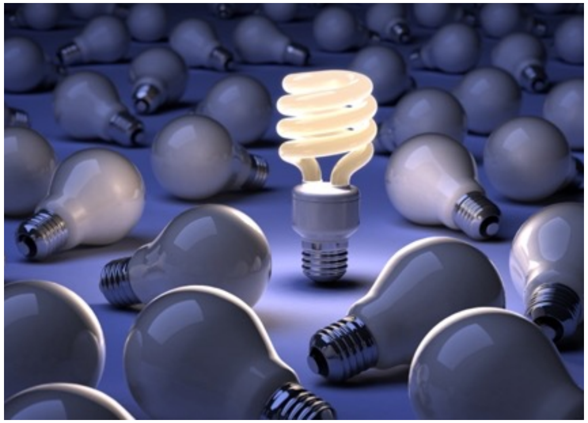


- Several challenges ahead
  - 99.97% reliable, but power outages still cost $150 billion/year
  - Customer engagement and environmental concerns

Picture source: FERC. "Final Report on the Aug. 14, 2003 Blackout in the US and Canada"

Features of Smart Grids

controllable



efficient

resilient

green/sustainable

self-healing

situational awareness

US Dept. of Energy, “The smart grid: An introduction.”

Enabling Technology Advances

renewables

power electronics

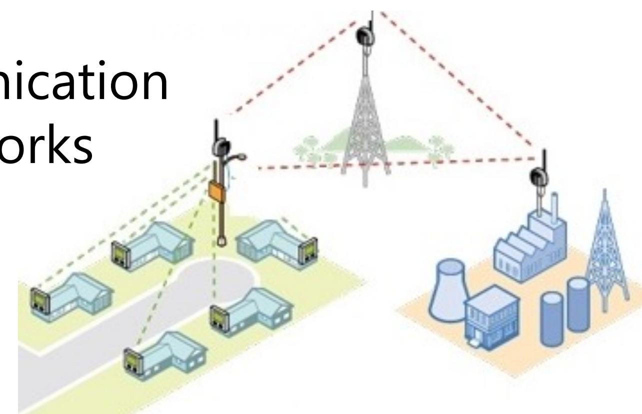
energy
storage

Optimization, signal processing, and machine learning toolbox

electric vehicles

metering

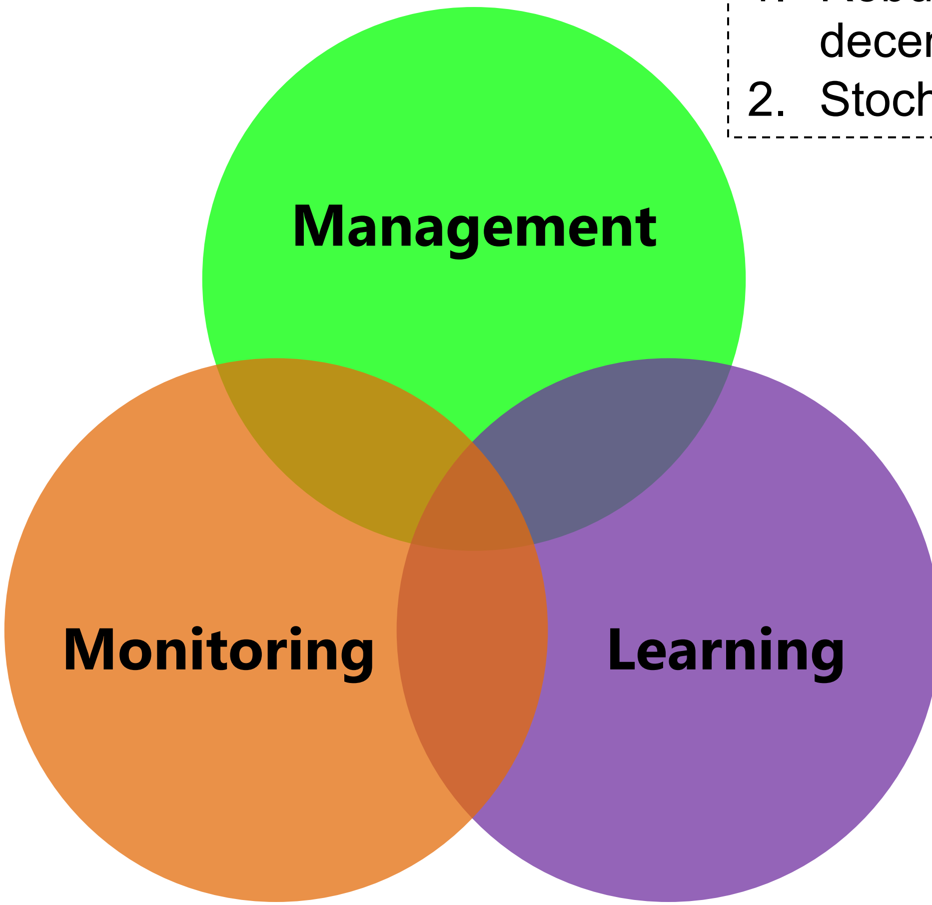
communication
networks



demand response

Research Thrusts

1. Robust/Stochastic microgrid decentralized power dispatch
2. Stochastic market clearing

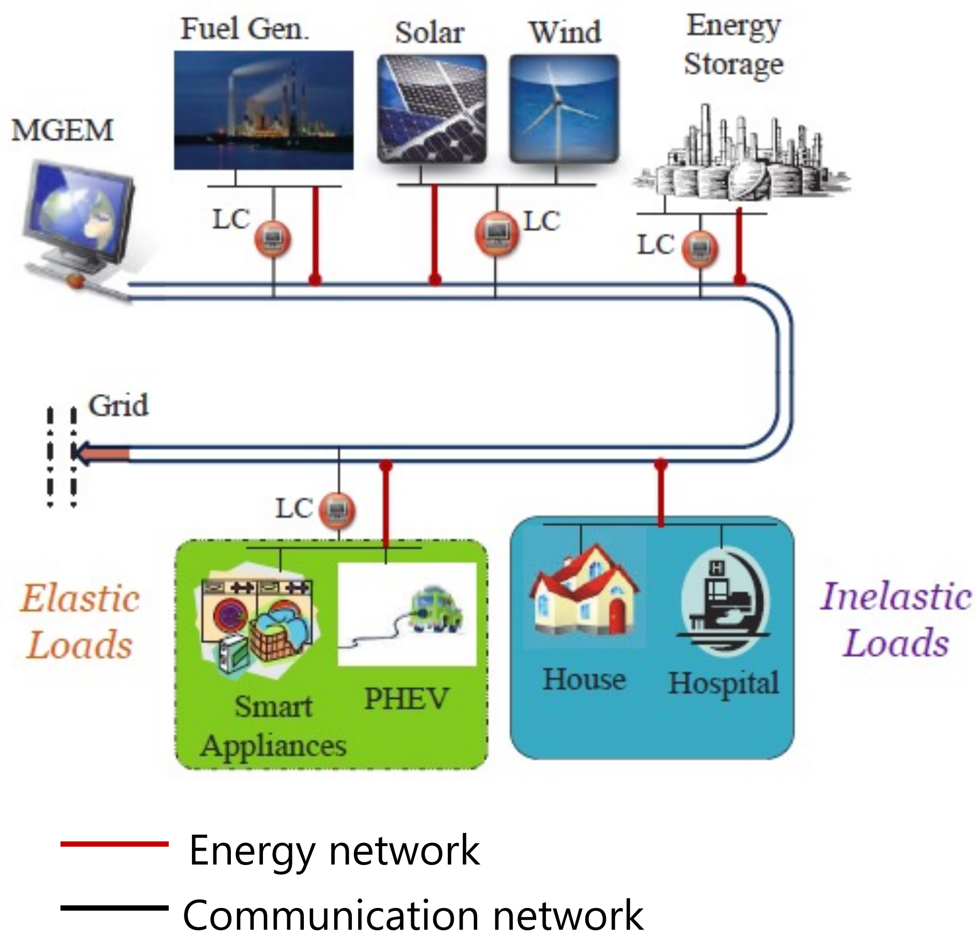


1. Efficient conic relaxations for power flow analysis and power system state estimation

1. Attention-based load forecasting
2. Dictionary learning for wind prediction
3. Multi-kernel learning for LMP forecasting
4. Fault location & NILM via deep learning
5. Semi-supervised event classification

Distributed Robust and Stochastic Energy Management

Micrograms and Distributed Energy Resources



Microgrids

- Distributed generation units
  - Fossil fuels
  - Renewables
- Distributed storage (DS)
- Elastic and inelastic loads

Challenges

- Renewable energy sources: high volatility
- Distributed scheduling over the microgrid infrastructure

Energy Management with Renewables

- Economic dispatch

$$\min_{\{P_{G_m}\}} \sum_{m=1}^{M} C_m(P_{G_m})$$
$$\text{s.t.} \quad \text{power generation constraints}$$
$$\sum_{m=1}^{M} P_{G_m} = P_D$$

(supply = demand)

$$\min_{\{P_{G_m}\}} \sum_{m=1}^{M} C_m(P_{G_m})$$
$$\text{s.t.} \quad \text{power generation constraints}$$
$$\sum_{m=1}^{M} P_{G_m} + W = P_D$$

stochastic resource

P_{G_m}: Power output of generator m

$C_m(\cdot)$: Cost of generator m

Problem Formulation

Goal: Minimize the microgrid net cost, subject to operational constraints

gen. cost | load utility | trans. cost

$$\min \left\{ \sum_{t=1}^{T}\sum_{m=1}^{M} C_m^t(P_{G_m}^t) - \sum_{t=1}^{T}\sum_{n=1}^{N} U_n^t(P_{D_n}^t) + G(\{P_R^t\}) \right\}$$

$$\text{s.t.}\quad P_{G_m}^{\min} \le P_{G_m}^t \le P_{G_m}^{\max},\ \forall\, m \in \mathcal{M},\ \forall\, t \in \mathcal{T}$$ generation limits

$$-R_m^{\text{down}} \le P_{G_m}^t - P_{G_m}^{t-1} \le R_m^{\text{up}},\ \forall\, m \in \mathcal{M},\ \forall\, t \in \mathcal{T}$$ ramping up/down

$$\sum_{m=1}^{M} (P_{G_m}^{\max} - P_{G_m}^t) \ge S^t,\ \forall\, t \in \mathcal{T}$$ spinning reserves

$$P_{D_n}^{\min} \le P_{D_n}^t \le P_{D_n}^{\max},\ \forall\, n \in \mathcal{N},\ \forall\, t \in \mathcal{T}$$ demand limits

$$P_R^{\min} \le P_R^t \le P_R^{\max},\ \forall\, t \in \mathcal{T}$$ committed power limits

$$\sum_{m=1}^{M} P_{G_m}^t + P_R^t = \sum_{n=1}^{N} P_{D_n}^t + L^t,\ \forall\, t \in \mathcal{T}$$ supply-demand balance

Key idea: use committed P_R^t , instead of actual W^t

Decision variables: $\{P_{G_m}^t, P_{D_n}^t, P_R^t\}$

Transaction Mechanism

ii) P_R^t : committed (to be delivered) energy at time t $\quad ([x]^+ := \max(x, 0))$

- shortage to buy: $[P_R^t - W^t]^+$ with purchase price α^t
- surplus to sell: $[W^t - P_R^t]^+$ with selling price β^t

❑ Worst-case transaction cost:

$$G(\{P_R^t\}) := \max_{\mathbf{w}\in\mathcal{W}} \left\{ \sum_{t=1}^{T} \left(\overbrace{\alpha^t [P_R^t - W^t]^+}^{\text{cost}} - \overbrace{\beta^t [W^t - P_R^t]^+}^{\text{revenue}} \right) \right\}$$

- protect the microgrid from suffering huge economic loss

Proposition 1 (convexity)

If the selling price β^t does not exceed the purchase price α^t for any $t \in \mathcal{T}$, then $G(\{P_R^t\})$ is convex in $\{P_R^t\}$.

Dual Decomposition

- Dual function (λ^t for balance eq., μ^t for spinning reserves)

$$\mathcal{D}(\{\mu^t\},\{\lambda^t\}) = \sum_{m=1}^{M} \min_{\{P_{G_m}^t\}} \left\{ \sum_{t=1}^{T} [C_m^t(P_{G_m}^t) + (\mu^t - \lambda^t) P_{G_m}^t] \right\} \quad \text{CG-LC}$$

$$+ \sum_{n=1}^{N} \min_{\{P_{D_n}^t\}} \left\{ \sum_{t=1}^{T} [\lambda^t P_{D_n}^t - U_n^t(P_{D_n}^t)] \right\} \quad \text{DSM-LC}$$

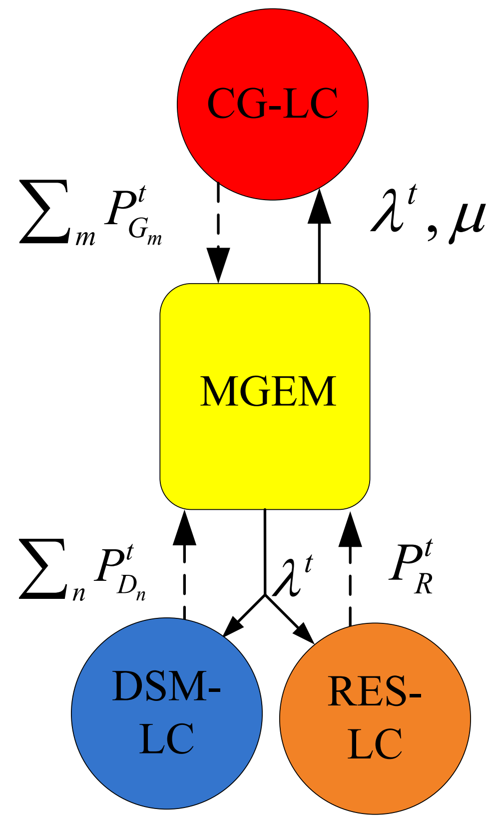
$$+ \min_{\{P_R^t\}} \left\{ G(\{P_R^t\}) - \sum_{t=1}^{T} \lambda^t P_R^t \right\} \quad \text{RES-LC}$$

CG-LC and DSM-LC: efficiently solvable: LP/QP

- RES-LC subproblem: proximal bundle method

- Dual update: subgradient projection

$$\mu^t(k+1) = [\mu^t(k) + a g_{\mu^t}(k)]^+$$
$$\lambda^t(k+1) = \lambda^t(k) + a g_{\lambda^t}(k)$$



Co-optimization of Renewables and Storage

- Joint uncertainty model for multiple wind farms

$$\mathcal{W} := \left\{ \mathbf{w} | \underline{W}_i^t \le W_i^t \le \overline{W}_i^t, W_s^{\min} \le \sum_{t \in \mathcal{T}_s} \sum_{i=1}^{I} W_i^t \le W_s^{\max}, \mathcal{T} = \bigcup_{s=1}^{S} \mathcal{T}_s \right\}$$

- Electric vehicles: $\sum_{t=S_q}^{T_q} P_{E_q}^t = E_q,\ P_{E_q}^t \in [P_{E_q}^{\min}, P_{E_q}^{\max}]$

- Distributed storage: $\boxed{B_j^t = B_j^{t-1} + P_{B_j}^t}$ $\quad B_j^t \in [0, B_j^{\max}], \quad B_j^T \ge B_j^{\min}$

  charging limits: $P_{B_j}^t \in \left[\max\left\{ P_{B_j}^{\min}, -\eta_j^{\text{dis}} B_j^{t-1} \right\}, \min\left\{ P_{B_j}^{\max}, \eta_j^{\text{ch}} (B_j^{\max} - B_j^{t-1}) \right\} \right]$

- Incorporate storage into the transaction

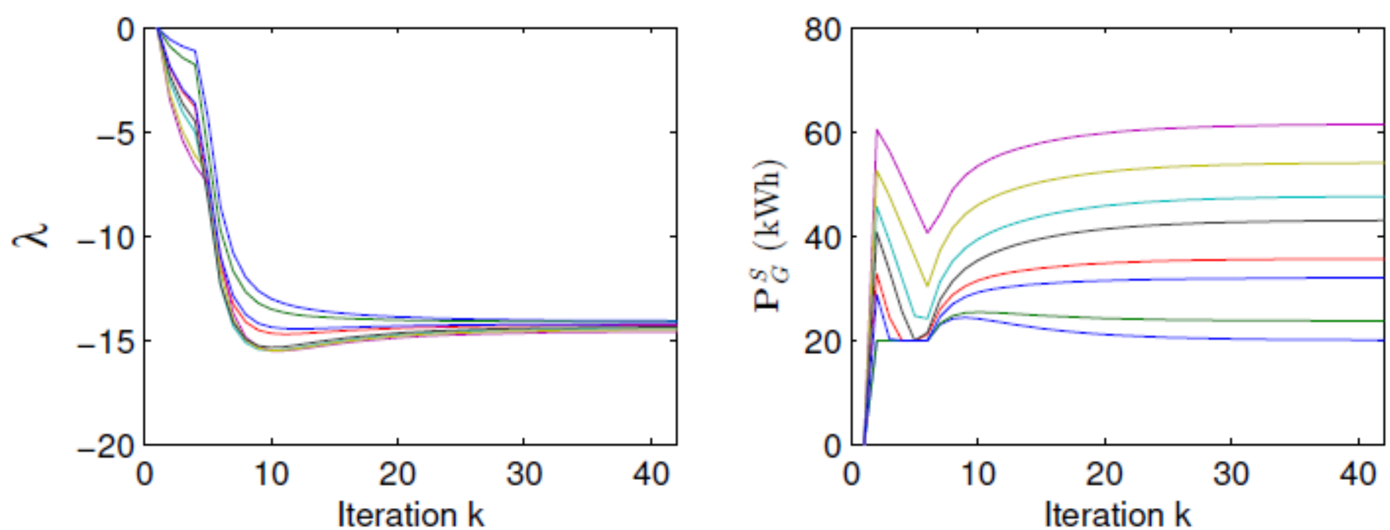
$$G(\{P_R^t\}, \{P_{B_j}^t\}) := \max_{\mathbf{w} \in \mathcal{W}} \sum_{t=1}^{T} \left(\alpha^t \left[P_R^t - \sum_{i=1}^{I} W_i^t + \sum_{j=1}^{J} P_{B_j}^t \right]^+ - \beta^t \left[\sum_{i=1}^{I} W_i^t - P_R^t - \sum_{j=1}^{J} P_{B_j}^t \right]^+ \right)$$

Stochastic Energy Management

❑ Consider the expected transaction cost

$$G(\mathbf{p}_R) := \mathbb{E}_{\mathbf{v}}\Big[\sum_{t=1}^{T}\Big(\alpha^t[P_R^t - \sum_{i=1}^{I} W_i^t(v_i^t)]^+ - \beta^t[P_R^t - \sum_{i=1}^{I} W_i^t(v_i^t)]^-\Big)\Big]$$

❑ Sample average approximation and ADMM



Yu Zhang and Georgios Giannakis, "Efficient decentralized economic dispatch for microgrids with wind power integration," *IEEE Green Tech.*, 2014.

More Works

- Yu Zhang and Georgios Giannakis, "Distributed Stochastic Market Clearing with High-Penetration Wind Power," *IEEE Trans. on Power Systems*, 2016.

- Gabriela Martinez, Yu Zhang, and Georgios Giannakis, "An Efficient Primal-Dual Approach to Chance-Constrained Economic Dispatch," *North American Power Symp. (NAPS)*, 2014.

- Yu Zhang, Nikolaos Gatsis, and Georgios Giannakis, "Risk-Constrained Energy Management with Multiple Wind Farms," *IEEE-PES on Innovative Smart Grid Tech. (ISGT)*, 2013.

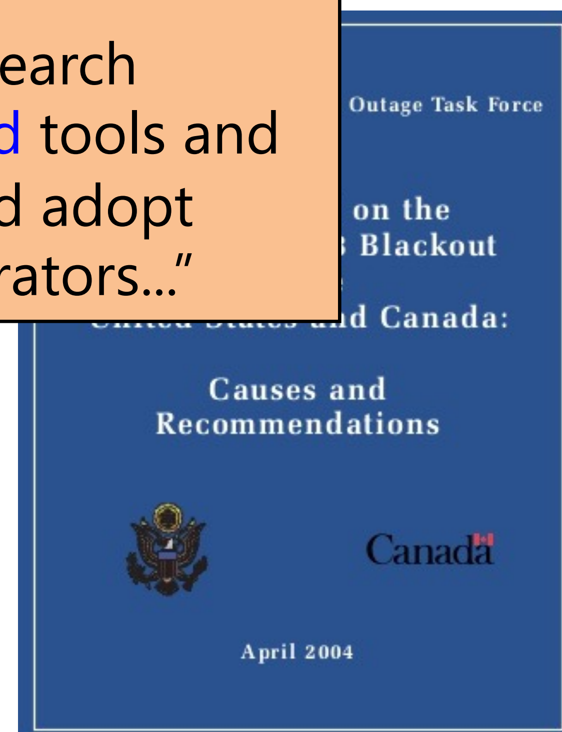
- Yu Zhang, Nikolaos Gatsis, Vassilis Kekatos, and Georgios Giannakis, "Risk-aware Management of Distributed Energy Resources," *Intl. Conf. on Digital Signal Process. (DSP)*, 2013.

State Estimation via Convex Relaxations

Remember …

Time: August 14, 2003
Location: Midwest/Northeast US & Ontario, CAN
Costs: 50 million people, 61,800 MWs of load lost. $4~10 billion in the US and Canada's GDP was down 0.7%

Recommendations:
1. "DOE should expand its research programs on reliability-related tools and technologies." 2. "Evaluate and adopt better real-time tools for operators…"

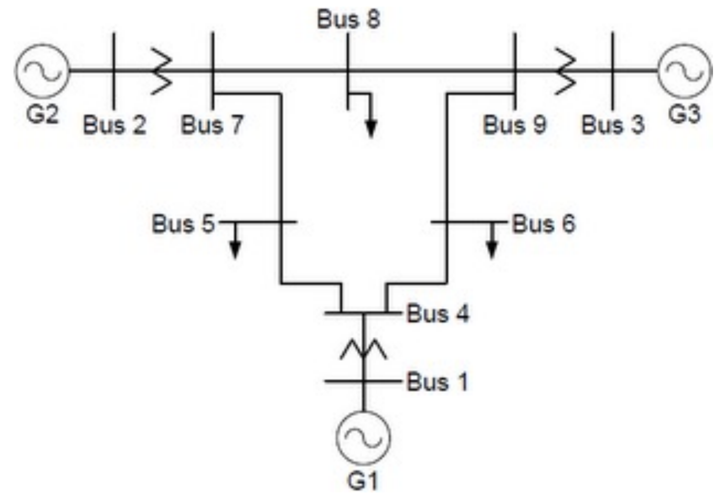


"**Key phase events:** MISO's state estimator (SE) software solution was compromised…" "The failure of its SE contributed to the lack of situational awareness."

System Modeling

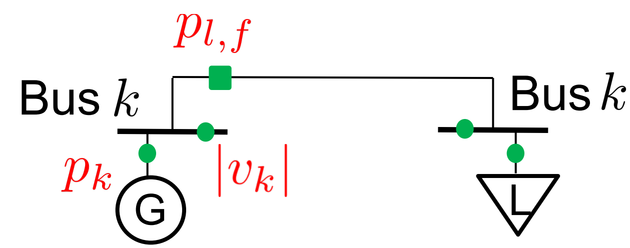
❑ A power system $\mathcal{G} = (\mathcal{N}, \mathcal{L})$

- Transmission lines, buses, and transformers



- Complex voltage: $\mathbf{v} = [v_1, \ldots, v_n]^\top \in \mathbb{C}^n$
- Nodal current injection: $\mathbf{i} = \mathbf{Y}\mathbf{v}$
- Net injected complex power: $\mathbf{p} + \mathbf{q}\mathrm{j} = \mathrm{diag}(\mathbf{v}\mathbf{i}^*)$

Nodal and Line Quantities



- Voltage magnitude and nodal power injections:

$$|v_k|^2 = \mathrm{Tr}(\mathbf{E}_k \mathbf{v}\mathbf{v}^*), \;\; p_k = \mathrm{Tr}(\mathbf{Y}_{k,p} \mathbf{v}\mathbf{v}^*), \;\; q_k = \mathrm{Tr}(\mathbf{Y}_{k,q} \mathbf{v}\mathbf{v}^*)$$

- Branch active and reactive powers:

$$p_{l,f} = \mathrm{Tr}(\mathbf{Y}_{l,p_f} \mathbf{v}\mathbf{v}^*), \quad p_{l,t} = \mathrm{Tr}(\mathbf{Y}_{l,p_t} \mathbf{v}\mathbf{v}^*)$$
$$q_{l,f} = \mathrm{Tr}(\mathbf{Y}_{l,q_f} \mathbf{v}\mathbf{v}^*), \quad q_{l,t} = \mathrm{Tr}(\mathbf{Y}_{l,q_t} \mathbf{v}\mathbf{v}^*)$$

- All quantities are quadratic functions of the complex $\mathbf{v}$ voltage

$$\mathbf{v} = \text{state of the system}$$

Problem Statement

Power system state estimation (PSSE):

Given noisy measurements $z_j = \mathrm{Tr}(\mathbf{M}_j \mathbf{v}\mathbf{v}^*) + \eta_j,\ j = 1, 2, \ldots, m,$ *estimate the complex voltage* $\mathbf{v}$.

❑ Functionality of PSSE:

- Provides real-time power system conditions
- Constitutes the core of online security analysis
- Provides diagnostics for modeling and maintenance

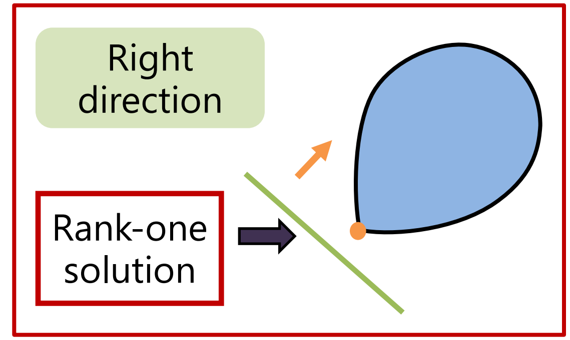
❑ Measurements:

- From SCADA system & phasor measurement units (PMUs)
- Challenges: grossly inaccurate due to noise and bad data
- Device malfunction, communication failure, and cyber attacks

Semidefinite Relaxation

- Our approach: Design a linear objective $\operatorname{Tr}(\mathbf{M}_0\mathbf{X})$ with $\mathbf{X} := \mathbf{v}\mathbf{v}^*$

$$
\begin{aligned}
\underset{\mathbf{X}\in\mathbb{H}^n}{\text{minimize}} \quad & \operatorname{Tr}(\mathbf{M}_0\mathbf{X}) \\
\text{subject to} \quad & \operatorname{Tr}(\mathbf{M}_j\mathbf{X}) = z_j,\ j \in \mathcal{M} \\
& \mathbf{X} \succeq \mathbf{0},\ \ \cancel{\operatorname{rank}(\mathbf{X}) = 1}
\end{aligned}
$$



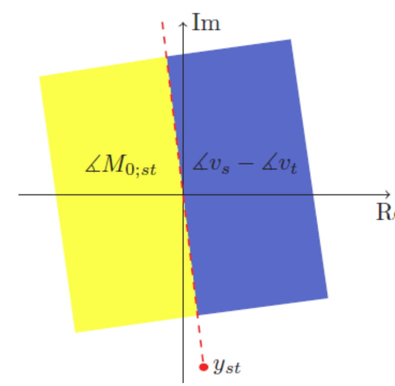
Question: *When is the SDP relaxation exact to recover* $\mathbf{v}$ *?*

- Assumptions

A1) Available measurements: $\begin{cases} |v_k|^2, \forall k \in \mathcal{N} \\ p_{l,f}\ \ (p_{l,t}), \forall l \in \mathcal{L}_{\mathrm{ST}} \end{cases}$

A2) Angle conditions:

$$
\begin{aligned}
-180^\circ < \measuredangle M_{0;st} - \measuredangle y_{st} < 0, &\quad \forall (s,t) \in \mathcal{L}_{\mathrm{ST}}, \\
0 < (\measuredangle v_s - \measuredangle v_t) - \measuredangle y_{st} < 180^\circ, &\quad \forall (s,t) \in \mathcal{L}_{\mathrm{ST}}, \\
(\measuredangle v_s - \measuredangle v_t) - \measuredangle M_{0;st} \neq 0 \ \text{or}\ 180^\circ, &\quad \forall (s,t) \in \mathcal{L}_{\mathrm{ST}}
\end{aligned}
$$



Power System State Estimation

- Penalized SDP:

$$\begin{array}{ll} \underset{\mathbf{X}\in\mathbb{H}^n,\boldsymbol{\nu}\in\mathbb{R}^m}{\text{minimize}} & \rho f(\boldsymbol{\nu}) + \mathrm{Tr}(\mathbf{M}_0\mathbf{X}) \\ \text{subject to} & \mathrm{Tr}(\mathbf{M}_j\mathbf{X}) + \nu_j = z_j, \quad \forall j \in \mathcal{M} \\ & \mathbf{X} \succeq \mathbf{0} \end{array}$$

- For example: $f_{\mathrm{WLAV}}(\boldsymbol{\nu}) = \sum_{j=1}^{m} |\nu_j|/\sigma_j, \quad f_{\mathrm{WLS}}(\boldsymbol{\nu}) = \sum_{j=1}^{m} \nu_j^2/\sigma_j^2$

Theorem 3 (performance bound)

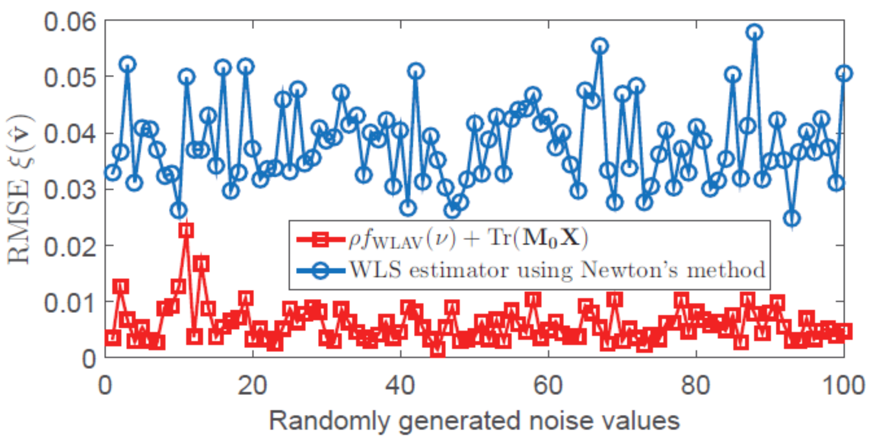
If $\rho \geq \max_{j\in\mathcal{M}} |\sigma_j \hat{\mu}_j|$, *then there exists a scalar* $\beta > 0$ *such that*

$$\zeta := \frac{\|\mathbf{X}^{\mathrm{opt}} - \beta\mathbf{v}\mathbf{v}^*\|_F}{\sqrt{n \times \mathrm{Tr}(\mathbf{X}^{\mathrm{opt}})}} \leq 2\sqrt{\frac{\rho \times f_{\mathrm{WLAV}}(\boldsymbol{\eta})}{n\lambda}},$$

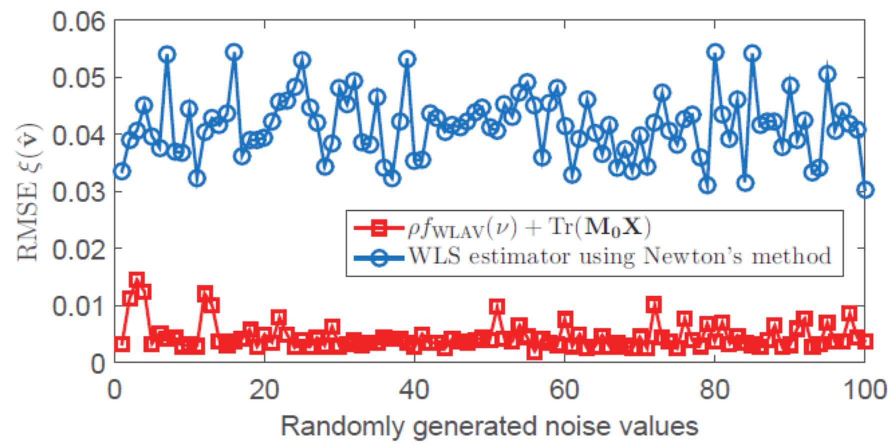
where λ *is the second smallest eigenvalue of* $\mathbf{H}(\hat{\boldsymbol{\mu}})$.

Penalized SDP vis-à-vis Newton's Method

- Performance metric: $\mathrm{RMSE} = \|\hat{\mathbf{v}} - \mathbf{v}_o\|/\sqrt{n}$
  - Available measurements: $\{|v_k|^2\}_{k\in\mathcal{N}}$ and $\{p_{l,f}, p_{l,t}\}_{l\in\mathcal{L}}$
  - Bad data: 20% of the line measurements



(a) IEEE 57-bus system



(b) IEEE 118-bus system

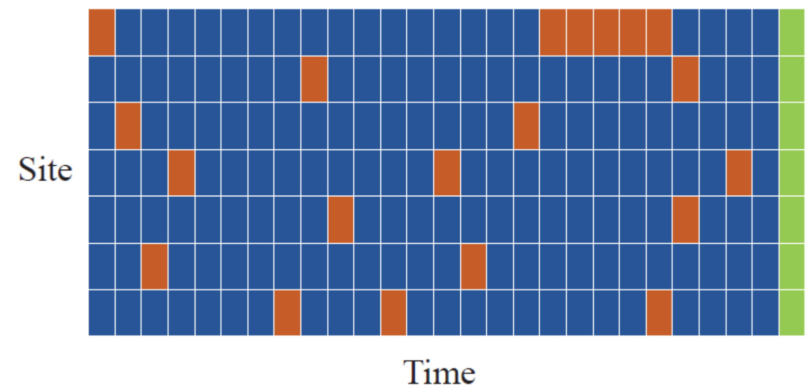
Energy Data Analytics

Energy Data Analytics

Figure source: B. P. Bhattarai *et al*., “Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions,” in *IET Smart Grid*, 2019.

Wind Power Prediction via Dictionary Learning

Wind power inference:
Given historical wind power outputs, infer missing and future values.



- Nonnegative dictionary learning and sparse coding

$$\min_{\mathbf{D}\in\mathcal{D},\mathbf{S}\in\mathcal{S}} \frac{1}{T}\sum_{t=1}^{T}\left\{\underbrace{\frac{1}{2}\|\mathbf{y}_t^{\text{obs}}-\mathbf{O}_t\mathbf{D}\mathbf{s}_t\|_2^2}_{\text{least-squares fitting error}} + \underbrace{\lambda_1\|\mathbf{s}_t\|_1+\frac{\lambda_2}{2}\|\mathbf{s}_t\|_2^2}_{\text{elastic-net regularizer}} + \underbrace{\frac{\lambda_L}{2}\mathbf{s}_t'\mathbf{D}'\mathbf{L}\mathbf{D}\mathbf{s}_t}_{\text{Laplacian regularizer}}\right\}$$

- Batch and online algorithms based on block coordinate descent

Yu Zhang, Seung-Jun Kim, and Georgios Giannakis, "Short-Term Wind Power Forecasting using Nonnegative Sparse Coding," *CISS*, 2015.

Low-rank Multi-kernel Learning for LMP Prediction

Market inference:

Assume that price $p(n,t)$ *at node n and hour t, depends on nodal features* $\mathbf{x}_n$ *and time features* $\mathbf{y}_t$.

Given historical feature-price pairs $\{(\mathbf{x}_n, \mathbf{y}_t), p(n,t)\}_{n\in\mathcal{N}, t\in\mathcal{T}}$,

infer prices $p(n',t')$ *at given* $\{(\mathbf{x}_{n'}, \mathbf{y}_{t'})\}_{n'\in\mathcal{N}', t'\in\mathcal{T}'}$.

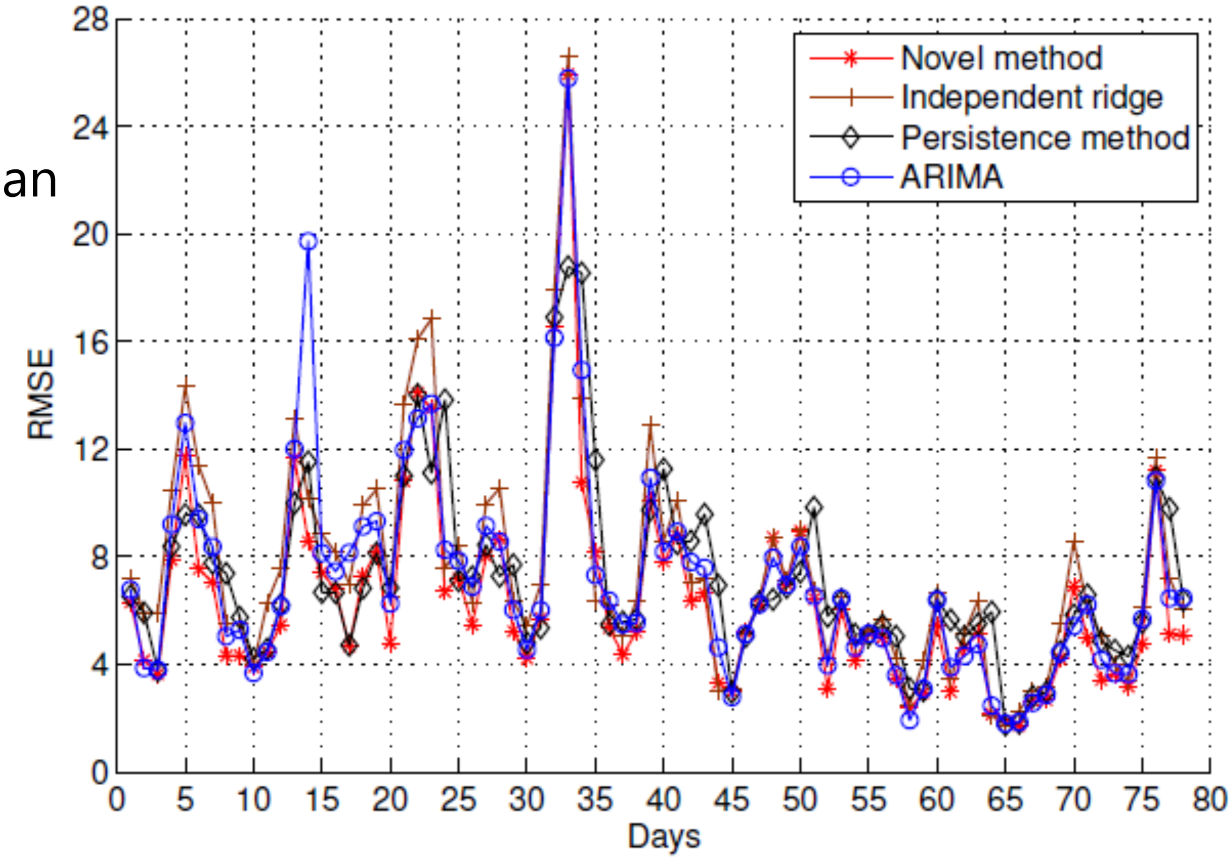
$$\min_{\mathbf{P},\{\mathbf{B}_l\},\{\mathbf{\Gamma}_m\}} \|\mathbf{Z}-\mathbf{P}\|_F^2 + \mu\sum_{l=1}^{L}\|\mathbf{B}_l\|_{\mathbf{K}_l} + \mu\sum_{m=1}^{M}\|\mathbf{\Gamma}_m\|_{\mathbf{G}_m}$$

$$\text{s.to} \quad \mathbf{P} = \sum_{l=1}^{L}\sum_{m=1}^{M}\mathbf{K}_l\mathbf{B}_l\mathbf{\Gamma}_m^\top\mathbf{G}_m.$$

- Nodal features
  - location
  - bus type: generator/load/interface
- Time features
  - yesterday's prices
  - temperature
  - day, hour, holiday

Day-ahead Prediction

- Nodal kernel
  - regularized Laplacian
  - diffusion Laplacian
  - Gaussian
  - identity
  - covariance
- Time kernel
  - Gaussian
  - linear



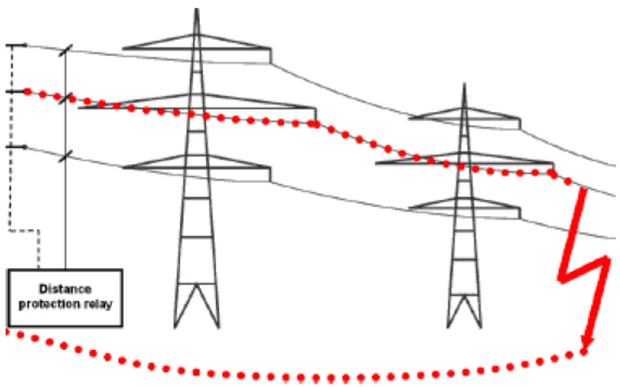
Average forecast error over MISO market (2012 summer)

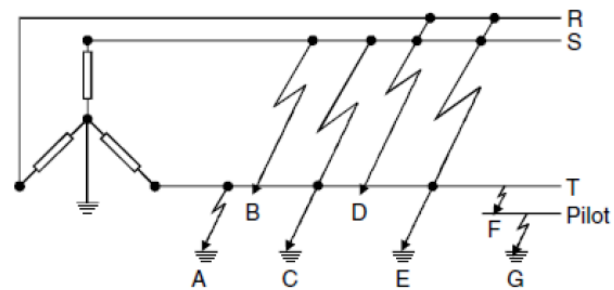
	Novel	Ridge	Persistence	ARIMA
RMSE	6.395	7.550	7.197	7.062
MAE	3.514	4.395	3.810	3.798

Vassilis Kekatos, Yu Zhang, and Georgios Giannakis, "Electricity Market Forecasting via Low-Rank Multi-Kernel Learning," *IEEE J. Selected Topics on Signal Proc.*, 2014.

Fault Location in Distribution Networks

- Power distribution systems are constantly under the threat of short-circuit faults that would cause power outages.





- System operators have to deal with outages timely to achieve high reliability.
- Goal: Accurately locate faults after the occurrence, so that quick restoration can be achieved.

Motivation

Existing methods:

- Methods based on impedance, voltage sag, traveling wave, and classical machine learning algorithms.
- They work fine theoretically, but are easily affected by noise, missing data, topology changes, etc.

The starting point of this work:

- Measuring phasors of voltage and current at lots of nodes becomes possible.
- Find a proper machine learning model that can use measurements from plenty of sensors to locate the fault in a distribution grid.

Problem Statement

We formulate the task of fault location as a classification problem.

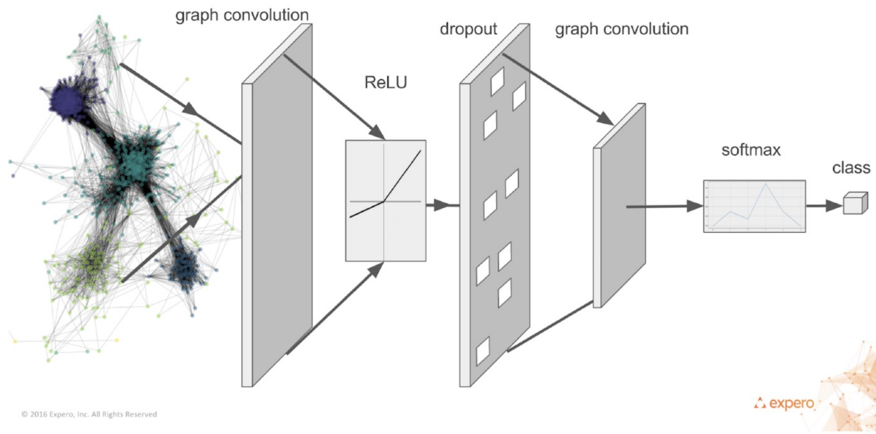
- For a given measured bus, assume that we have access to its three-phase voltage and current phasors:

$$(V_1, \theta_1^V, V_2, \theta_2^V, V_3, \theta_3^V, I_1, \theta_1^I, I_2, \theta_2^I, I_3, \theta_3^I) \in \mathbb{R}^{12}.$$

- A data sample of measurements: $\mathbf{X} \in \mathbb{R}^{n_o \times 12}$, where n_o is the number of observed buses. Values for unmeasured phases are set to zero.
- Given a data sample $\mathbf{X}_i$, the faulty bus $\tilde{y}_i = f(\mathbf{X}_i)$, where f is a classification model. A fault is correctly located if $\tilde{y}_i = y_i$, where y_i indicates the true faulty bus.

K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He. "Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks." *IEEE J. on Selected Areas in Commu.*, 2019.

Graph Convolutional Networks (GCN)



- Graph is a natural representation of a power network: nodes for buses; edges for power lines.
- GCNs is a powerful predictive tool: leverages the information contained in the data and the relationships between data.

- Chebyshev polynomial filters stabilizes the training with low computational complexity:

$$\mathbf{\Phi} h_{\boldsymbol{\alpha}}(\mathbf{\Lambda})\mathbf{\Phi}^{\top}\mathbf{f} = h_{\boldsymbol{\alpha}}(\mathbf{L})\mathbf{f} = \sum_{k=0}^{K} \alpha_k T_k(\tilde{\mathbf{L}})\mathbf{f}$$

K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He. "Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks." *IEEE J. on Selected Areas in Communications*, 2019.

Calculate the Laplacian

- Find the distance matrix $\mathbf{S} \in \mathbb{R}^{n\times n}$. The entry $\mathbf{S}_{ij}$ is the length of the shortest path between bus i and bus j.

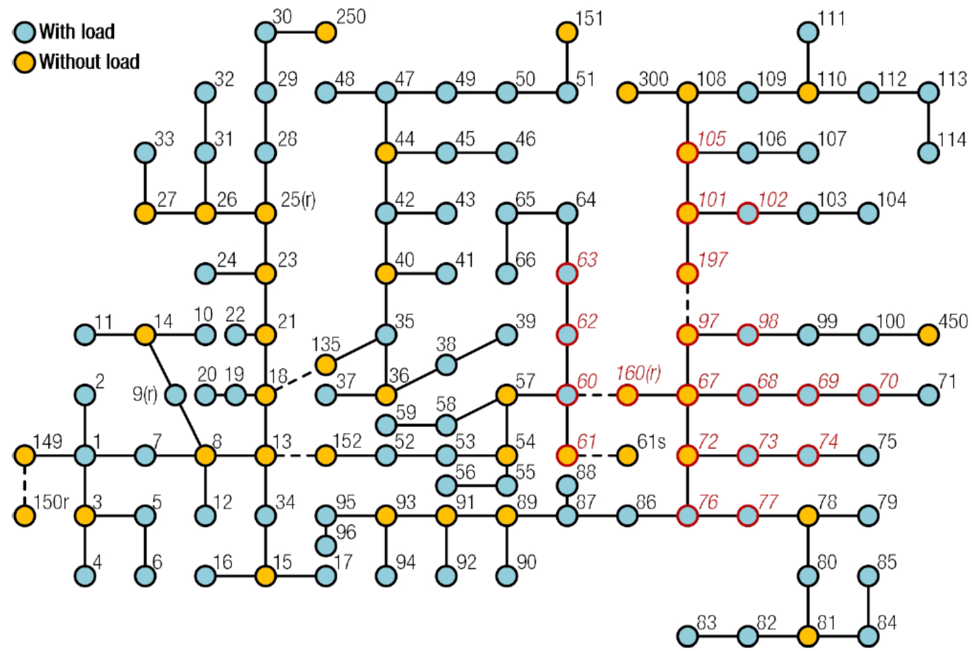
- Ascending-sort (from left to right) and keep the smallest K_n values in each row of $\mathbf{S}$ to obtain $\tilde{\mathbf{S}} \in \mathbb{R}^{n\times K_n}$ and calculate the normalization factor $\sigma_S = \sum_i \tilde{\mathbf{S}}_{iK_n}/n$.

- Calculate $\tilde{\mathbf{W}}_{ij} = e^{-\tilde{\mathbf{S}}_{ij}^2/\sigma_S^2}$. Obtain the weighted adjacency matrix $\mathbf{W} \in \mathbb{R}^{n\times n}$ by restoring the positional correspondence of $\tilde{\mathbf{W}}_{ij}$ to bus i and bus j.

- Calculate the Laplacian matrix $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$.

IEEE 123-Bus Test Case

Three types of faults:
- 1 phase to ground
- 2 phases to ground
- 2 phase short-circuit



Model	Noise	Noise + Bus	Noise + Random	All Combined
PCA + SVM	85.70 / 96.21	55.98 / 77.74	58.00 / 80.17	44.12 / 68.51
PCA + RF	86.51 / 97.55	64.11 / 81.61	66.12 / 84.90	52.34 / 72.13
FCNN	86.95 / 97.19	61.95 / 82.58	70.32 / 88.52	53.98 / 76.55
GCN	**97.52 / 99.73**	**92.67 / 98.26**	**88.76 / 96.44**	**84.53 / 94.77**

Energy Disaggregation via CNN

- **Energy disaggregation** (a.k.a. nonintrusive load monitoring, NILM): Decompose the whole energy consumption of a house into the energy usage of individual appliances (single-channel blind source separation problem).

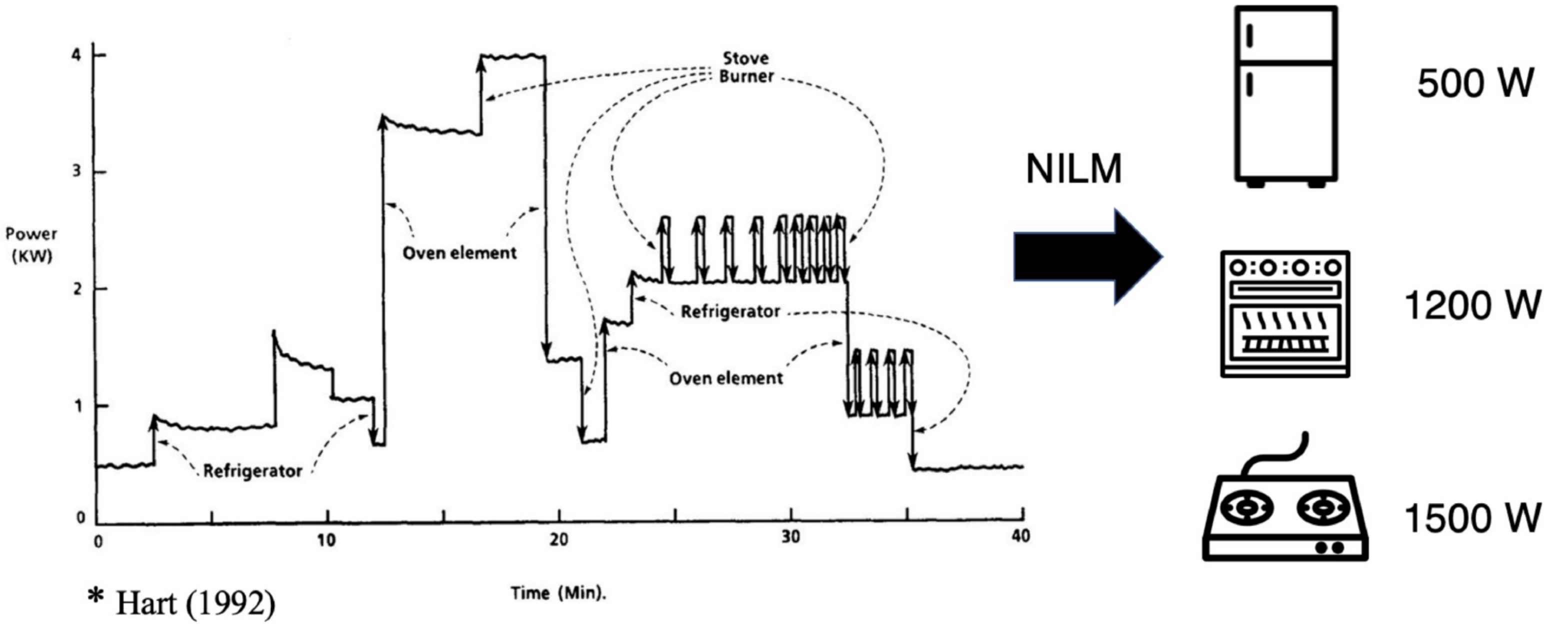


Figure: github.com/ch-shin/awesome-nilm

Prior Work

Models and algorithms for energy disaggregation:

- **Supervised**: k-means and SVM [Altrabalsi et al. (2014)], various neural network models including RNN, CNN, and denoising auto-encoders [Kelly and Knottenbelt (2015)].
- **Unsupervised**: factorial hidden Markov model [Kolter and Jaakkola (2012)], graph signal processing [Zhao et al. (2016)].
- **Other methods**: appliance-wise rules based on state changes [Farinaccio and Zmeureanu (1999)].

We propose the scale- and context-aware network (SCANet) in this work.
Incorporate the understanding of the task's challenges into the model:

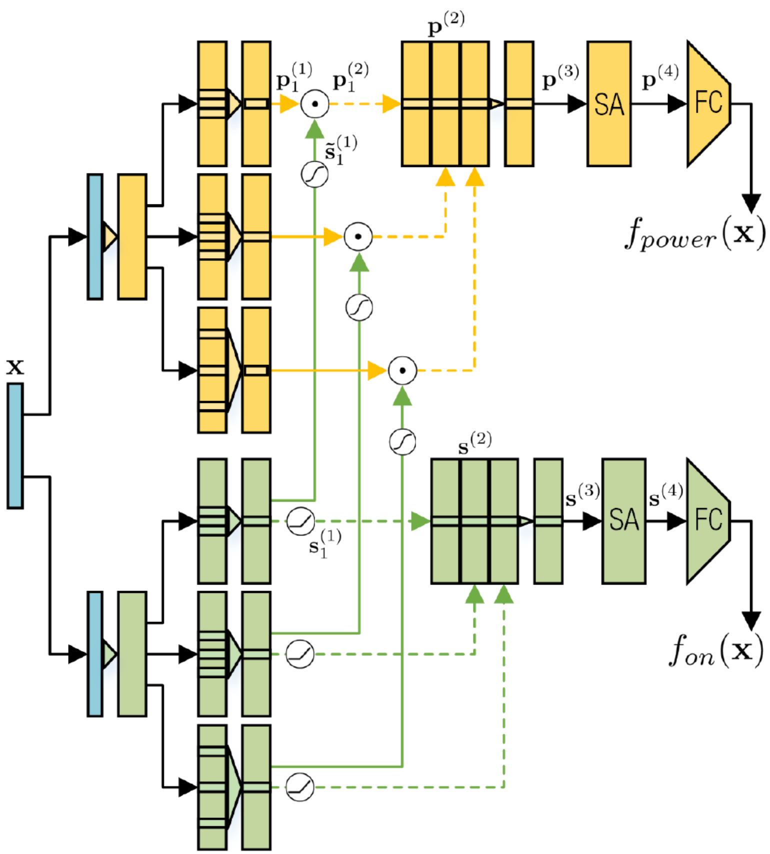
- Energy consumption behaviors contain higher-level *domain knowledge*: e.g., i) dryer works after washer; and ii) microwave can be turned on multiple times until cooking is finished.
- Time scale variance of different appliances: Requiring the ability to deal with *scale variation*.

K Chen, Y Zhang, Q Wang, J Hu, H Fan, J He. "Scale- and Context-Aware Convolutional Non-intrusive Load Monitoring," *IEEE Transactions on Power Systems*, 2020.

New Contributions

How to independently accomplish the tasks of **regression** & **classification**?

- Estimating the power consumed by appliance i at time t: $\hat{\mathbf{p}}_t^i \in \mathbb{R}_+^s$
- Identifying the on/off state of appliance i at time t: $\hat{\mathbf{o}}_t^i \in [0, 1]^s$.



1. Scale-awareness: dilated convolution
2. Attention mechanism: extract context info
3. On-state augmentation: for peak power of fridge
4. Critic: facilitate more realistic outputs

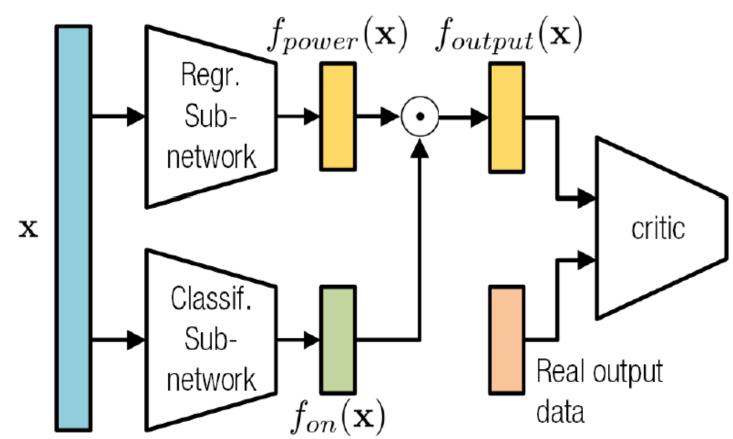


Figure: The structure of the proposed SCANet.

Estimation Errors

The REDD dataset

Improvement: 22.6% for MAE and 28.2% for SAE

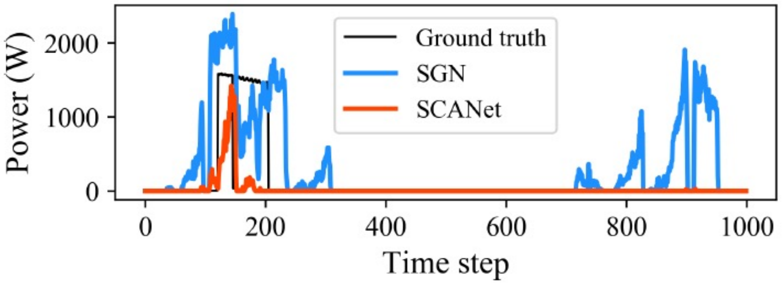
Metric	Model	Fridge	Microwave	Dishwasher	Average
MAE	Seq2point [Zhang *et al.*, 2018a]	26.01	27.13	24.15	25.76
	SGN [Shin *et al.*, 2018]	26.11	16.95	15.88	19.65
	Proposed SCANet	**21.77**	**13.75**	**10.12**	**15.21**
SAE	Seq2point	16.24	18.89	22.98	19.37
	SGN	17.28	12.49	14.59	14.79
	Proposed SCANet	**14.05**	**9.97**	**7.83**	**10.62**

The UK-DALE dataset

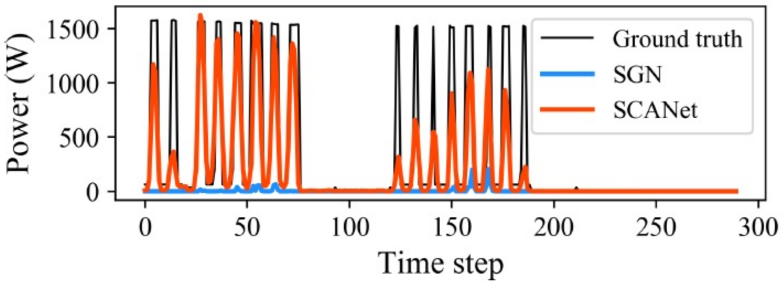
Improvement: 9.3% for MAE and 15.5% for SAE

Metric	Model	Washing machine	Kettle	Fridge	Microwave	Dishwasher	Average
MAE	Seq2point	15.09	11.43	18.01	17.87	18.50	16.18
	SGN	10.40	8.22	14.66	6.52	11.11	10.18
	Proposed SCANet	**8.84**	**7.36**	**14.37**	**5.53**	**10.06**	**9.23**
SAE	Seq2point	11.98	5.60	**6.99**	14.62	12.35	10.31
	SGN	7.99	5.35	9.24	5.11	8.43	7.22
	Proposed SCANet	**6.63**	**4.32**	8.50	**4.34**	**6.71**	**6.10**

Results for Microwave



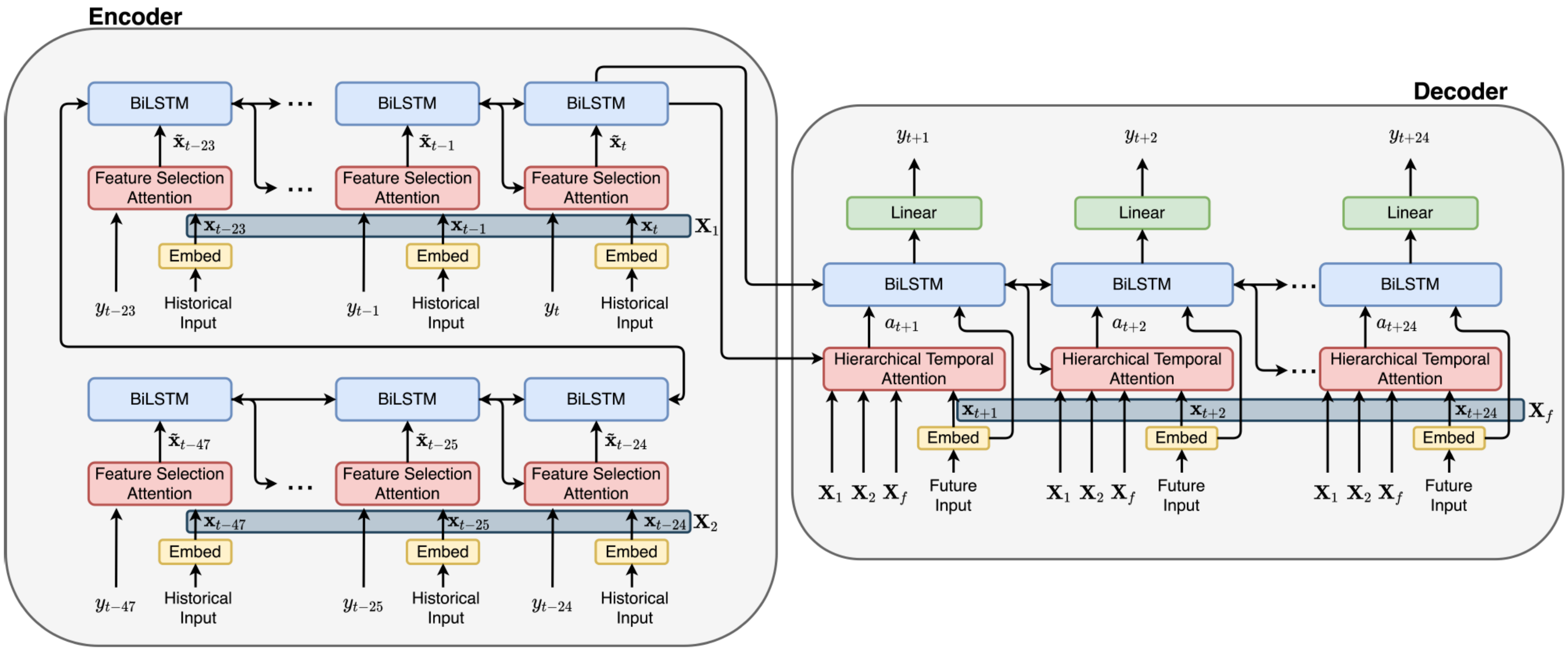
(a) False positive example of SGN



(b) A case SCANet is able to detect

- (a) SGN tends to have more false positive sections.
- (b) SCANet can identify that the microwave consumes power for multiple successive short durations.

Attention-based Load Forecasting



- The feature selection attention in the encoder is designed to adaptively weigh the different input features.
- The hierarchical temporal attention in the decoder focuses more on the temporal similarity to incorporate similar day information.

J. Xiong, P. Zhou, A. Chen, and Y. Zhang. "Attention-based Neural Load Forecasting: A Dynamic Feature Selection Approach." *Accepted by IEEE PES-GM 2021.*

Forecasting Errors

Model	MAE	RMSE	MAPE (%)	nRMSE (%)
GBM	71.48	94.61	2.15	2.84
RF	85.36	113.67	2.55	3.42
SVR	79.23	108.27	2.33	3.26
EDLSM	74.73	97.95	2.26	2.94
EDBiLSTM	71.48	92.94	2.17	2.79
ANLF	**64.80**	**86.74**	**1.93**	**2.61**

Online Detection of Faults & Cyber Attacks

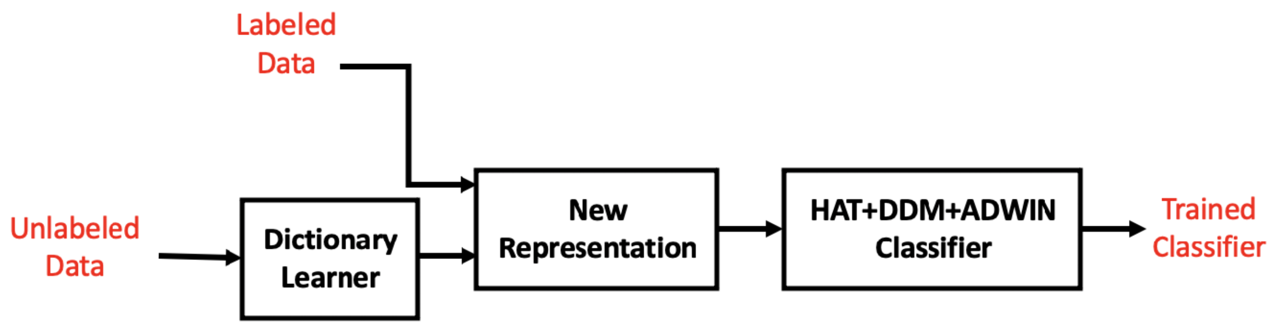


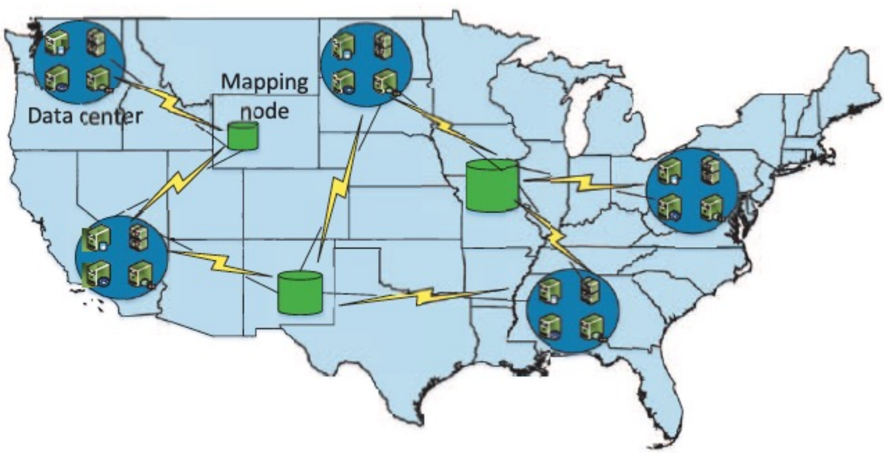
Figure: Diagram of our proposed algorithm, named as SSHAD (semi-supervised HAT+DDM+ADWIN).

We extend Adhikari's work by leverageing semi-supervised learning:

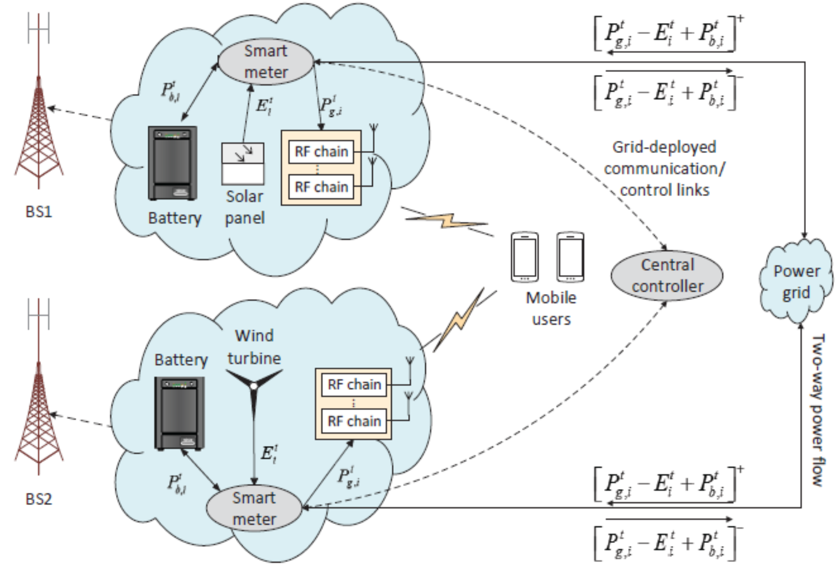
- First, we learn higher-level features of the unlabeled data, and store it on a dictionary.
- Second, we build a new feature representation of the labeled dataset, using the dictionary. The orginal labels are preserved.
- Third, the new labeled dataset is fed into a HAT+DDM+ADWIN classifier. As a result, we have the trained classifier.

G. Intriago and Y. Zhang. "Semi-Supervised Online Detection of Cyber Attacks and Faults for Power Systems." *Accepted by IEEE PES-GM 2021.*

Renewable Powered Cyber-Physical Systems



Geo-distributed data centers



Cellular networks

System features:

- Locally supported by renewables and storage units
- Two-way energy trading

Challenges:

- Uncertainties from renewables, prices, and service requests
- Distributed resource allocation