

Essi Purhonen

**Hybridimallit tuulen nopeuden monivaiheisessa
ennustamisessa**

Tietotekniikan Kandidaatintutkielma

17. kesäkuuta 2024

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Essi Purhonen

Yhteystiedot: `essi.e.purhonen@student.jyu.fi`

Ohjaaja: Leevi Annala

Työn nimi: Hybridimallit tuulen nopeuden monivaiheisessa ennustamisessa

Title in English: Hybrid models in multi-step wind speed forecasting

Työ: Kandidaatintutkielma

Sivumäärä: 45+0

Tiivistelmä: Tässä kandidaatintutkielmassa perehdytään tuulen nopeuden monivaiheisessa ennustamisessa käytettyihin hybridimalleihin ja menetelmiin, jotka vaikuttavat positiivisesti mallien suorituskykyyn. Tutkielman johtopäätöksenä on, että hybridimallit, jotka hyödyntävät datan esikäsitteilyä, datan jälkikäsitteilyä ja optimointimenetelmiä saavuttavat parhaimpia tuloksia.

Avainsanat: tekoäly, ennustaminen, tuulen nopeus, optimointialgoritmit, hajottaminen, ominaisuuksien valinta, ominaisuuksien irrottaminen,

Abstract: This bachelor's thesis focuses on hybrid models used in multi-step wind speed prediction and on methods that have a positive impact on the model's performance. The conclusion of this thesis is hybrid models that use data preprocessing, data postprocessing and optimization methods have the most accurate results.

Keywords: artificial intelligence, prediction, wind speed, optimization, decomposition, feature selection, feature extraction

Termiluettelo

CEEMD	Täydennetty ensemble-empiirinen moodin hajottaminen. Muunnelma EMD:stä ja EEMD:stä.
ICEEMDAN	Paranneltu täydennetty ensemble-empiirinen moodin hajoaminen, josa adaptiivista kohinaa. Muunnelma EMD:stä ja CEEMDAN:sta.
CEEMDAN	Täydennetty ensemble-empiirinen moodin hajoaminen, josa adaptiivista kohinaa. Muunnelma EMD:stä.
CSA	Crow-etsintäalgoritmi. Optimointialgoritmi.
EEMD	Ensemble-empiirinen moodin hajottaminen. Muunnelma EMD:stä.
EMD	Empiirinen moodin hajottaminen. Menetelmä, jolla aikasarjojen analyysiin.
EWT	Empiirinen aallokemuunnos. Itsemukautuva datan hajottamisen menetelmä.
IEWT	Käänteinen empiirinen aallokemuunnos. Suodatinpohjainen datan jälkikäsitteilyn menetelmä.
IHWOSCA	Improved Hybrid Grey Wolf Optimizer Sine Cosine Algorithm. Optimointialgoritmi.
IMF	Sisäinen moodifunktio. Hajottaessa sarja EMD-menetelmällä syntyy joukko näitä stationaarisempia komponentteja.
MOWOA	Multi-Objective Whale Optimization Algorithm. WOA:n monitavoiteoptimointiin soveltuva muunnos.
OVMD	Optimaalinen variaatiomoodin hajottaminen. Itseadaptoituva signaalin käsittelymenetelmä, joka juontuu VMD:stä.
PSO	Hiukkaspervioptimointi. Optimointialgoritmi.
SDA	Hybridi tai sekundaarinen harjoittamisen algoritmi, jossa yhdistetään useampaa hajottamisen menetelmää.
VMD	Variaatiomoodin hajottaminen. Signaalin käsittelymenetelmä.
WOA	Whale Optimization Algorithm. Optimointialgoritmi, joka perustuu ryhävalaiden saalistuskäyttäytymiseen.

WPF	Wavelet Packet Filter. Suodatinpohjainen datan jälkikäsittelyn menetelmä.
WPD	Wavelet Packet Decomposition. Signaalin käsittelymenetelmä, joka hajottaa aikasarjan. Muunnelma WT:stä.
WT	Aallokemuunnos. Signaalin käsittelymenetelmä, joka hajottaa aikasarjan.

Kuviot

Kuvio 1. Koonti hybridimallien suorituskyvyn arvioinneista	28
--	----

Taulukot

Taulukko 1. Ennustushorisontit, niiden aikavälit ja applikaatiot (Soman ym. 2010).	5
Taulukko 2. Tuulen nopeuden ennustustulokset datajoukoille A-C (Du ym. 2017).	19
Taulukko 3. Mittarit, joilla arvioidaan mallien suorituskykyä. (Niu ym. 2018)	21
Taulukko 4. MAPE arvojen keskimääräinen parantuminen (%) kaikille datajoukoille. (Sun ym. 2018).....	23
Taulukko 5. TVF-EMD-FE-SSA-MHHOGWO-PSR-KELM-ConvLSTM mallin ennustustulosten MAPE(%) arvot. (Fu ym. 2020).....	24
Taulukko 6. EWT-GWO-RELM-IEWT-mallin suorituskyvyn arviointi. (Liu, Wu ja Li 2018)	25
Taulukko 7. Hybridimallin suorituskyvyn keskimääräinen arviointi kaikille datajou- koille. (Liu ja Chen 2019b)	27

Sisällys

1	JOHDANTO	1
2	TUULEN ENNUSTAMINEN	3
2.1	Hyvin lyhyt johdatus aikasarjoihin	3
2.2	Tuulen ennustamisen tekniikat	4
2.2.1	Ennustushorisontit	5
2.2.2	Ennustamismenetelmät	6
3	HYBRIDIMALLIT	8
4	DATAN KÄSITTELYMENETELMÄT HYBRIDIMALLEISSA	9
4.1	Datan esikäsittelyn menetelmiä	10
4.1.1	Datan hajottamisen menetelmät	10
4.1.2	Ominaisuuksien valintaa ja ominaisuuksien irrottamista hyödyntävät mallit	14
4.1.3	Kohinan poistaminen	15
4.1.4	Poikkeavien havaintojen havaitseminen	15
4.2	Datan jälkikäsittelyn menetelmät hybridimalleissa	16
5	OPTIMOINTIMENETELMÄT TUULEN NOPEUDEN ENNUSTAMISESSA	17
5.1	Optimointi hybridimalleissa	17
5.2	Monitavoiteoptimointi hybridimalleissa	18
6	HYBRIDIMALLIT TUULEN NOPEUDEN MONIVAIHEISESSA ENNUSTAMISESSA	21
6.1	Mallien suorituskyvyn mittarit	21
6.2	Datan esikäsittelyä ja optimointia hyödyntävät hybridimallit	21
6.3	Datan esikäsittelyä, jälkikäsittelyä ja optimointia hyödyntävät hybridimallit ..	24
7	YHTEENVETO	29
	LÄHTEET	31

1 Johdanto

Ilmastonmuutos ja kasvihuonepäästöjen kasvu ovat herättäneet tutkijoiden kiinnostuksen uusiutuviin energialähteisiin (Yang ym. 2019). Eurooppa on määritellyt pyrkivänsä hiilineutraaliksi vuoteen 2050 mennessä ja Suomi vuoteen 2035 mennessä (Valtioneuvosto 2019). Nämä tavoitteet kasvattavat tarvetta puhtaiden ja ympäristöystävällisten energiamuotojen lisäämiseksi energiantuotannossa. Uusiutuvista energialähteistä tuulienergian määrä kasvaa nopeinten sen laajan saatavuuden, hiilineutraaliuuden ja taloudellisuuden ansiosta (Dhiman, Deb ja Balas 2020a). Tuulen epävakaa ja epäsäännöllinen luonne tekee tuulienergiasta haasteellisen lisättävän sähköverkkoon (Dhiman, Deb ja Balas 2020a; C. Li ym. 2018). Tuulen ennustamista on tutkittu paljon viimeisen parin vuosikymmen aikana. Kuitenkin tuulen ennustaminen on edelleen haasteellista, ja tarkemmille ennusteille on tarve. Esimerkiksi sähköverkon vakauden ja luotettavuuden takaamiseksi tuulen nopeuden ennustaminen on tärkeää (Fu, Wang, Li ym. 2019). Lyhytaikaisen (eng. short-term) tuulen nopeuden ennustaminen on myös oleellista kuormituksen lähettämisessä ja energiamarkkinoiden selvityksessä (Dhiman, Deb ja Balas 2020a). Tuulen nopeuden ennustamisessa käytettävät menetelmät voidaan jakaa fysikaalisiin, klassisiin tilastotieteellisiin, tekoäly- ja hybridimalleihin. Näistä hybridimallit ovat osoittautuneet tarkimmiksi tuulen nopeuden ennustamisessa (Mosavi ym. 2019). Lisäksi hybridimallit, joissa hyödynnetään datan esikäsittelyä ja optimointityökaluja näyttävät pystyvän saavuttamaan vielä tarkempia tuloksia.

Tämän tutkimuksen tarkoitus on kartoittaa koneoppimis pohjaisia hybridimalleja lyhytaikaisessa tuulen nopeuden ennustamisessa ja vertailla mallien tarkkuutta. Tutkielmaan valitut hybridimallit ovat monivaiheisia (eng. multi-step) malleja. Monivaiheisista hybridimalleista on aineistoa, mutta kokoavia katselmuksia ei löydy. Koska teoriakenttä tarvitsee jotain kokoavaa, halusin valita monivaiheiset hybridimallit tutkielmaani. Lisäksi monivaiheisilla malleilla on suurempi merkitys käytännön sovelluksissa. Tutkimus on suoritettu kirjallisuuskatsauksena painottaen uutta tutkimustietoa ja artikkeleiden viittausten määrää.

Luvussa 2 käsitellään tuulen ennustamisen paradigmoja ja luokittelua. Luku 3 käsittelee hybridimalleja. Luvussa 4 esitellään hybridimalleissa käytettyjä datan käsittelymenetelmiä. Luku 5 käsittelee optimointialgoritmeja hybridimalleissa. Luvussa 6 tarkastellaan viittä eri hy-

bridimallia tarkemmin. Ja luvussa 7 on tutkielman pääkohtien yhteenveto ja pohditaan tuulen ennustamisen tulevaisuutta.

2 Tuulen ennustaminen

Ilmastonmuutoksen hillitsemiseksi ja ilmaston lämpenemisen rajaamiseksi 1.5 celsiusasteeseen Pariisin ilmastonsopimuksen mukaisesti (Secretariat 2016) tarve ympäristöystävällisille ja hiilineutraaleille energialähteille on suuri. IPCC tuo esille raportissaan (Masson-Delmotte ym. 2018), että pystyäksemme pitäytymään 1.5 celsiusasteen lämpötilan nousussa tulee 70-85% energiantuotannosta toteutua uusiutuvilla energialähteillä. Tuuli on yksi vanhimpia käytettyjä energialähteitä. Se on myös yksi suosituimpia uusituvista energialähteistä sen taloudellisuuden, puhtauden ja laajan saatavuuden ansiosta (Dhiman, Deb ja Balas 2020a). Tuulen epälineaarisen ja epästabiilin luonteen vuoksi sen ennustaminen ja sähköverkkoon lisääminen on haasteellista (Dhiman, Deb ja Balas 2020a; C. Li ym. 2018). Sähköverkon vakauden ja luotettavuuden takaamiseksi tuulen nopeuden ennustaminen on tärkeää, sillä se oleellisesti vaikuttaa tuulivoimaan (Fu, Wang, Li ym. 2019). Tuulivoima lasketaan yhtälöstä

$$P = \frac{1}{2} \rho A v^3, \quad (2.1)$$

missä ρ on ilman tiheys (kg/m^3), A on roottorin ala (m^2) ja v on tuulen nopeus (m/s) (Nazir ym. 2020). Kaavasta (2.1) nähdään, että tuulen nopeuden ja voiman välinen suhde on kuutiollinen, jonka seurauksena pieni virhe tuulen nopeuden ennustuksessa aiheuttaa suuren virheen ennustettavassa tuulivoimassa. Lisäksi tuulen nopeuteen ja tuulivoimaan vaikuttavat useat tekijät, kuten ilman lämpötila, ilmanpaine, kosteus ja tuulimyllyjen sijainti voimalassa ja voimalaa ympäröivä maasto (Dhiman, Deb ja Balas 2020b). Koska tuulen nopeuden ennustaminen on tuulen nopeuden aikasarjojen ennustamista, seuraavassa kappaleessa käsitellään lyhyesti aikasarjoja.

2.1 Hyvin lyhyt johdatus aikasarjoihin

Tuulienergian ennustaminen voidaan nähdään stokastisena prosessina ja tuulen nopeuden muutoksia suhteessa aikaan mallinnetaan aikasarjojen avulla (Liu, Chen ym. 2019). Dhiman, Deb ja Balas (2020b) määrittelee aikasarjan “peräkkäisinä järjestettyinä sarjoina numeerisia datapisteitä ja se voidaan ottaa mistä tahansa ajallisesti muuttuvasta muuttujasta”. Aikasarja-analyysi mahdollistaa muuttujan käytöksen tarkastelun suhteessa ajanhetkestä

seuraavaan. Aikasarjat voidaan jakaa stationaarisiin ja ei-stationaarisiin. Stationaarinen aikasarja s_t voidaan ilmaista seuraavasti

$$F_S(s_{t_1+h}, \dots, s_{t_n+h}) = F_S(s_{t_1}, \dots, s_{t_n}) \quad (2.2)$$

missä t on aika-askel ja h on eteenpäin siirtynyt aika-askel (Dhiman, Deb ja Balas 2020b). Eli aikasarja on stationaarinen, jos sen tilastolliset parametrit, kuten varianssi, ei ole ajasta riippuvia. Tilastollisin testein, esimerkiksi autokorrelaatiofunktio (eng. autocorrelation function, ACF) ja Ljung-Box-testin avulla voidaan määrittää onko sarja stationaarinen vai ei (Dhiman, Deb ja Balas 2020b). Tuulen aikasarja on hyvin ei-stationaarinen (Qian ym. 2019).

2.2 Tuulen ennustamisen tekniikat

Tuulivoiman ja nopeuden ennustamisen tarkoituksena on antaa informaatiota sähkövoimajärjestelmille tuulen käytöksestä tarvittavalle ajanjaksolle. Kirjallisuudessa nousee esille kaksi yleistä luokittelua tuulen nopeuden ennustamiselle: ennustamismenetelmä ja ennustamisen aikahorisontti (Liu ja Chen 2019a; Dhiman, Deb ja Balas 2020b; Nazir ym. 2020; Ahmadi ja Khashei 2021; J. Wang ym. 2016). Aikahorisonttien mukaiselle jaottelulle on tarve erityisesti tuulivoiman ja tuulen nopeuden ennustamisen käytännön sovelluksilla, kuten sähkömarkkinoiden selvityksellä (eng. clearing) ja kuormituksen lähetyksellä (Dhiman, Deb ja Balas 2020a). Ennustamismallien tarkoituksena on välittää tietoa tuulen nopeudesta ja voimasta tulevan aikavälin, aikahorisontin, verran. Aikahorisontin pituus vaikuttaa oleellisesti, mihin tarkoitukseen kyseistä mallia käytetään (Liu ja Chen 2019a).

Kirjallisuudessa on nähtävissä yksivaiheisia (eng. single-step) (Memarzadeh ja Keynia 2020), ja monivaiheisia (eng. multi-step) (J. Wang ym. 2016) ennustuksia. Monivaiheisilla ennustuksilla on merkitys käytännön sovelluksissa, koska ne antava energiajärjestelmille aikaa operoida sähköverkkoja esimerkiksi poikkeustilanteissa (Y. Li ym. 2019). Wang ja Hu (2015) määrittelevät monivaiheisen ennustamisen seuraavasti: olkoon t ennustuksen ajanhetken alkupiste ja olkoon $h, h > 0$ ennustushorisontti. Olkaamme kiinnostuneita ennustamaan piste v'_{t+h} . Olkoon $v'_t(h)$ pisteen v_{t+h} ennuste. Tällöin $v'_t(h)$ on pisteen v_x ennustus h -askeleen päähän ajanhetken alkupisteestä t . Jos $h = 1$, niin puhutaan yhden askeleen (eng. one-step) ennustuksesta.

Tämä kirjallisuuskatsaus fokusoituu monivaiheisiin ennustuksiin, koska niillä on suurempi merkitys käytännön sovelluksissa. Tämä kappale käsittelee seuraavaksi tuulen ennustamisen aikahorisontteja ja päättyy eri ennustamismenetelmien esittelyyn.

2.2.1 Ennustushorisontit

Ennustamisen tekniikat voidaan jakaa aikahorisonttien mukaan. Kirjallisuudessa puhutaan sekä ennustushorisontista (eng. forecasting tai prediction horizon) (Liu ja Chen 2019a) sekä aikahorisontista (eng. time horizon) (Ahmadi ja Khashei 2021). Nämä tarkoittavat samaa asiaa eli ajan pituutta, jota ennustus koskee. Ennustushorisonteilla on erityinen merkitys käytännön applikaatioissa, ja ne helpottavat oikean menetelmän valintaa käyttökohteen perusteella (Dhiman, Deb ja Balas 2020b). Ennustushorisontit voidaan kategorisoida hyvin lyhyen aikavälin (sekunneista noin 30 minuuttiin), lyhyen aikavälin (30 minuutista kuuteen tuntiin), keskipitkän aikavälin (kuudesta tunnista päivään) ja pitkän aikavälin (päivästä viikkoon tai pidemmäksi ajaksi) ennustuksiksi. Jaottelu ei ole tarkkaa (Liu ja Chen 2019a). Taulukossa 1 on eräs kuvaus ennustushorisonteista, aikaväleistä ja applikaatioista.

Taulukko 1. Ennustushorisontit, niiden aikavälit ja applikaatiot (Soman ym. 2010).

Ennustushorisontti	Aikaväli	Applikaatio
hyvin lyhyt	sekunneista	- sähkömarkkinoiden selvitys
	30 minuuttiin	- säätelytoimet
lyhyt	30 minuutista	- kuormituksen lähetys
	6 tuntiin	- kuormituksen kasvatus/vähennys
keskipitkä	6 tunnista	- generaattorin päälle/pois päältä
	24 tuntiin	kytkentöjen päätökset - käyttövarmuus (päivää edeltävä) sähkömarkkinoilla
pitkä	24 tunnista viikkoon	- yksikön sitoutumispäätökset
	tai pidemmälle	- varantovelvoitteiden päätökset -ylläpidon aikataulutuksen optimaalisten käyttökustannuksien saavuttamiseksi

2.2.2 Ennustamismenetelmät

Tuulen nopeuden ennustaminen voidaan jakaa seuraaviin neljään kategoriaan 1. fysikaalisiin, 2. tilastotieteellisiin, 3. älykkäisiin ja 4. hybridimalleihin (H. Liu ym. 2020). Jokaisella mallilla on omat vahvuutensa ja soveltuvuutensa eri aikahorisontin mukaisiin ennustuksiin. Kuitenkin kirjallisuus viittaa, että hybridimallit ovat soveltuvimpia uusiutuvien energialähteiden, kuten tuulen, ennustamisessa (Mosavi ym. 2019).

1. Fysikaalisten menetelmien mallit hyödyntävät meteorologista dataa ja tuulen nopeuden syitä, kuten orografisuutta (vuoristot), maanpinnan muotoja, ilmanpainetta, korkeutta ja lämpötilaa (Niu ja Wang 2019; Liu, Chen ym. 2019). Menetelmiä ovat mm. numeerinen sään ennustaminen (eng. numerical weather prediction, NWP) (Cassola ja Burlando 2012) ja sään tutkimus- ja ennustamismalli (eng. weather research and forecastinf model, WRFM) (Zhao ym. 2016). Mallit soveltuvat hyvin pitkäkestoisen tuulen nopeuden ennustamiseen, mutta soveltuvat huonosti lyhytkestoiseen (Ahmadi ja Khashei 2021). Lisäksi mallit ovat laskennallisesti raskaita ja resursseja tuhlaavia, koska vaativat suuria määriä dataa (Jiang ja Liu 2019).

2. Tilastollisissa menetelmissä käytetään ennustamiseen suurta määrää historiallista dataa ja tilastollisia menetelmiä, kuten todennäköisyyslaskentaa ja stokastisia prosesseja (Liu, Chen ym. 2019). Yleisimpiin tilastotieteellisiin aikasarjamalleihin kuuluvat esimerkiksi autoregressiivinen liukuvan keskiarvon malli (eng. Auto-Regressive Moving Average, ARMA), autoregressiivinen integroitu liikkuvan keskiarvon malli (eng. Autoregressive Integrated Moving Average, ARIMA) ja murto-ARIMA (eng. fractional ARIMA) (Ahmadi ja Khashei 2021). (Kavasseri ja Seetharaman 2009) käyttivät murto-ARIMA mallia ennustamaan tuulen nopeutta 24 ja 48 tunnin aikahorisontissa. Tilastolliset menetelmät soveltuvat hyvin mallintamaan lineaarisia, yksinkertaisempi aikasarjoja, sillä ne perustuvat oletukseen aikasarjojen lineaarisuudesta. Tuulen nopeuden sarja on kuitenkin hyvin epälineaarinen. (Jiang ja Liu 2019). Tämän vuoksi ainoastaan tilastollisia menetelmiä hyödyntävät tuulen nopeuden ennustukset eivät ole täsmällisiä (eng. accuracy).

3. Tekoälymenetelmiä käyttävien mallien hyvä kyky ennustaa ei-lineaarisia aikasarjoja on lisännyt niiden käyttöä tuulen nopeuden ennustamisessa viime vuosina. Tekoälymenetel-

mien etuina pidetään esimerkiksi niiden adaptiivisuutta, virheiden sietokykyä (eng. fault tolerance) ja itseorganisoituvuutta (eng. self-organizing) (Marugán ym. 2018). Tekoälymallit ovat osoittautuneet nopeiksi, tehokkaiksi ja toimiviksi menetelmiksi ja ovat vakiintuneet tuulen nopeuden ennustamisessa (Mosavi ym. 2019). Tuulen nopeuden ennustamisessa käytettyjä koneoppimismalleja on useita, kuten erilaiset neuroverkot (eng. artificial neural network, ANN) (Liu, X. Mi ja Y. Li 2018b; Song, Wang ja Lu 2018), äärimmäiset oppimiskoneet (eng. extreme learning machine, ELM) (Liu, X. Mi ja Y. Li 2018a) ja tukiverkkokoneet (eng. support vector machines, SVM) (M. Liu ym. 2020; Zhou, Shi ja Li 2011). Syväoppimismenetelmien kehittyessä niiden käyttö on lisääntynyt tuulen ennustamisessa. Lisäksi siirtovaikutusoppiminen (eng. transfer learning) tuo mielenkiintoisia mahdollisuuksia tuulen nopeuden ja tuulivoiman ennustamiseen esimerkiksi lyhentämällä laskennallista aikaa ja parantamalla ennustamisen täsmällisyyttä (Liu, Chen ym. 2019). Vaikka älykkäät mallit ovat osoittautuneet toimiviksi metodeiksi tuulen nopeuden ennustamisessa, niilläkin on heikkouksensa. Esimerkiksi mallit helposti saattavat päätyä lokaaliin optimiin, konvergoitua hitaasti tai ylisovittua (Niu ja Wang 2019).

Kaikki ylläolevat mallit sisältävät puutteita ja eivät yksinään kykene täsmälliseen tuulen nopeuden ennustamiseen. Seuraava kappale käsittelee hybridimalleja, jotka on kehitetty vastaamaan yksittäisten mallien ongelmiin.

3 Hybridimallit

Hybridimalleihin liittyvää kirjallisuutta löytyy paljon ja kirjallisuuden perusteella ne soveltuvat tuulienergian ennustamiseen paremmin kuin yksittäiset mallit. Kirjallisuudessa ei ole tarkkaa määritelmää hybridimalleille. Mosavi ym. (2019) mukaan hybridimalleissa hyödynnetään useampaa koneoppimismallia ja/tai pehmeän laskennan metodeja. Jotkut liittävät kokonaisuusoppimisen (eng. ensemble learning) osaksi hybridimalleja, mutta toisaalta Kazienko, Lughofer ja Trawinski (2015) määrittelee hybridimenetelmät ja kokonaisuusoppimisen menetelmät erikseen. Tässä kartoituksessa määrittelen koneoppimis pohjaiset hybridimallit malleina, joissa yhdistetään koneoppimismenetelmään yhtä tai useampaa koneoppimisen ja/tai pehmeän laskennan menetelmiä.

Kirjallisuuden perusteella hybridimallit voidaan karkeasti jakaa datan prosessointia hyödyntäviin malleihin, optimointialgoritmeja hyödyntäviin malleihin ja komponentteja yhdistäviin malleihin (Ahmadi ja Khashei 2021; Liu ja Chen 2019a). Komponentteja yhdistelevissä ennustusmalleissa yhdistetään useampaa koneoppimisen menetelmää ja/tai tilastotieteellistä menetelmää, joiden erillisistä ennustuksista lopullinen ennustettu aikasarja muodostetaan. Datan prosessointi voidaan jakaa vielä datan esikäsittelyyn ja jälkikäsittelyyn. Datan käsittelytekniikat ja erilaiset optimointialgoritmit ovat parantaneet ennustusten täsmällisyyttä ja stabiiliutta (Ahmadi ja Khashei 2021). Nykyisin hybridimallit, jotka yhdistävät sekä datan käsittelyn että optimointialgoritmien käytön ovat yleistyneet lupaavien tuloksien ansioista. Seuraavissa kappaleissa esitellään hybridimalleissa käytettyjä datan käsittelytekniikoita ja optimointimenetelmiä.

4 Datan käsittelymenetelmät hybridimalleissa

Tuulen nopeuden ennustamisessa käytetyt tilastotieteelliset ja älykkäät mallit ovat hyvin datavetoisia. Erilaisia datan käsittelyyn tarkoitettuja menetelmiä on otettu käyttöön mallien ennustuksien täsmällisyyden parantamiseksi. Hybridimalleissa esiintyy sekä datan esikäsittelyä (Zhou, Liu ym. 2018) että jälkikäsittelyä (Liu, Wu ja Li 2018). Useat monivaiheiset hybridimallit hyödyntävät molempia tekniikoita (Wang ja Li 2018). Esikäsittelyllä pyritään saavuttamaan pienempi ennustusvirhe vähentämällä syötteen kompleksisuutta ja poistamalla kohinaa. Tämä myös alentaa laskennallista kuormitusta vähentämättä ennustuksen täsmällisyyttä. Jälkikäsittely taas parantaa alustavaa mallin ennustusta käsittelemällä tulostetta ja siinä esiintyvää virhettä (Liu ja Chen 2019a).

Hybridimalleissa esiintyviä datan käsittelytekniikoita Liu ja Chen (2019a) kategorisoi seuraavasti

- hajottaminen (eng. decomposition)
- ominaisuuksien valinta (eng. feature selection)
- ominaisuuksien irrotus (eng. feature extraction)
- kohinan poistaminen (eng. denoising)
- jäännösvirheen mallintaminen (eng. residual error modelling)
- poikkeavien havaintojen havaitseminen (eng. outlier detection)
- suodatinpohjainen korjaus (eng. filter-based correction).

Käsittelytekniikka valitaan tarvittavan käytön mukaisesti. Yleisesti hajottamista käytetään jakamaan alkuperäinen ei-stationaarinen aikasarja useampiin stationaarisiiin osasarjoihin, jotka ovat helpommin ennustettavissa. Ominaisuuksien valinta ja irrotus taas pienentää dimensiota ja siten keventää laskennallista taakkaa vähentämättä ennustamisen täsmällisyyttä. Kohinan ja poikkeavien havaintojen poistamisella esikäsitellään dataa ja siten parannetaan ennustuksen täsmällisyyttä. Jäännösvirheiden mallintamista sekä suodatinpohjaista korjausta hyödynnetään jälkikäsittelyssä (Liu ja Chen 2019a).

4.1 Datan esikäsittelyn menetelmiä

Tuulen nopeuden aikasarjan epästabiiliisuuden vuoksi yksittäisten mallien tuottamat ennustukset eivät ole täsmällisiä. Yksittäisen mallin on vaikea mallintaa alkuperäistä, kompleksista aikasarjaa (Qian ym. 2019). Datan esikäsittelyllä pystytään tätä alkuperäisen datan kompleksisuutta vähentämään ja siten saavuttamaan täsmällisempiä ennustuksia (Z. Liu ym. 2020). Datan esikäsittelyn hyödyntäminen tuulen ennustamisessa onkin kasvanut 28% vuosien 2015-2019 aikana (Ahmadi ja Khashei 2021). Kappaleessa 4 mainituista datan käsittelymenetelmistä hajottaminen, ominaisuuksien valinta ja irrotus sekä kohinan poisto ovat esikäsittelymenetelmiä. Tämä kappale esittelee lyhyesti aiemmin mainittujen esikäsittelytekniikoiden yleisimpiä algoritmeja ja menetelmiä.

4.1.1 Datan hajottamisen menetelmät

Datan hajottamista hyödyntävissä hybridimalleissa tarkoituksena on parantaa ennustusta hajottamalla alkuperäinen tuulen aikasarja useiksi osasarjoiksi, jotka ovat stabiilimpia ja siten helpommin mallinnettavissa kuin alkuperäinen volatiliteetti aikasarja. Jokaiselle osasarjalle muodostetaan oma ennustusmalli (Qian ym. 2019). Hajottamista hyödyntävät hybridimallit voidaan kategorisoida käytettävien hajottamisen algoritmien mukaan. Liu ja Chen (2019a) jaottelee mallit viiteen kategoriaan

1. aallokepohjaisiin (eng. wavelet-based)
2. empiirinen moodin hajoaminen eli EMD-pohjaisiin
3. variaatiomoodin hajoamiseen (eng. variational mode decomposition, VMD)
4. itsemukautuva hajottaminen (eng. self-adaptive decomposition)
5. hybrideihin eli sekundaarisiin hajottamisen algoritmeihin (eng. Secondary Decomposition Algorithm, SDA)

Qian ym. (2019) ei kategorisoi erikseen itsemukautuvia ja hybridejä, mutta edellä mainittujen kolmen kategorian lisäksi mainitsee myös kausitasoitusmenetelmän (eng. Seasonal Adjustment Method, SAM) ja Intrinsic Time-Scale Decomposition, ITD. Tämä tutkielma ei käsittele kahta viimeksi mainittua menetelmää, koska niitä ei hyödynnetä tutkielmaan valituissa hybridimalleissa, ja niitä ei esiintynyt muissa alan katsauksissa.

Kaikissa datan hajottamisen tekniikoissa on sama yleisidea, jossa alkuperäinen aikasarja hajotetaan stabiilimmiksi osasarjoiksi ennen uudelleen käsittelyä. Kuitenkin Qian ym. (2019) on havainnut kolme eri rakennetta hajottamista hyödyntäville malleille. Ensimmäisessä rakenteessa jokainen osasarja analysoidaan ja ennustetaan käyttäen erilaisia malleja, joista edelleen itsenäiset ennustukset yhdistetään lopulliseksi tulokseksi. Toisessa rakenteessa, joka muistuttaa Liu ja Chen (2019a) hybridejä eli sekundaarisia hajottamisen algoritmeja, osasarjat hajotetaan edelleen ja jatkokäsitellään ensimmäisen rakenteen mukaisesti. Kolmannessa rakenteessa osasarjoista ominaisuuksien valintaa hyödyntämällä valitaan syöte yhteen ennustusmalliin. Seuraavaksi esitellään lyhyesti yleisimmät hajottamisen tekniikat.

Aallokepohjainen hajottaminen

Aallokepohjaisessa hajottamisessa (eng. Wavelet-based decomposition) aikasarja hajotetaan alipäästö- ja ylipäästösuodattimilla approksimoiduksi sarjaksi ja yksityiskohtaisemmaksi sarjaksi (Qian ym. 2019). Qian ym. (2019) ja Liu ja Chen (2019a) viittaavat hieman eri termein aallokemuunnokseen. Liu ja Chen (2019a) puhuu aallokemuunnoksesta (eng. wavelet transform, WT) ja Qian ym. (2019) käyttää termiä aallokehajoittaminen (eng. wavelet decomposition, WD). Kirjoittajat tarkoittavat kuitenkin samaa asiaa. Käytän tässä WT eli aallokemuunnosta. WT hajottaa aikasarjan jokaisella binääripuun tasolla matalataajuisen komponentin ja lopulta jäljellä on joukko, jossa on yhden matalataajuisen approksimoidun osasarja ja useita korkeantaajuuden yksityiskohtaisia osasarjoja (Memarzadeh ja Keynia 2020).

Kuitenkin tuulen nopeuden aikasarja sisältää paljon informaatiota korkeammilla taajuuksilla, siksi on oleellista hajottaa myös korkean taajuuden yksityiskohtaiset komponentit (Liu ja Chen 2019a). Wavelet Packet Decomposition, WPD, on muunnelma WT:stä, jossa jokaisella tasolla hajotetaan sekä matalan taajuuden että korkean taajuuden komponentti ja saadaan joukko sekä matalan että korkean taajuuden osasarjoja (Y. Li ym. 2019). Näin saavutetaan kattavampi informaation jakautuminen (Liu ja Chen 2019a). WPD näyttää tuottavan parempia tuloksia kuin WT, koska WPD saa enemmän tietoa alkuperäisestä tuulen nopeuden datasta, mutta se on myös laskennallisesti raskaampi (Qian ym. 2019; Liu ja Chen 2019a). Vaikka aallokepohjainen hajottaminen toimii hyvin ei-stationaarisen datan, kuten tuulen nopeuden, analysoinnissa, se kuitenkin luonteeltaan hyvin joustamaton, ei-adaptiivinen, ja sen käyttö vaatii paljon ennakkotietoja ja suuren määrän kokeellista dataa saavuttaakseen hyviä

tuloksia käytännön sovelluksissa (Liu ja Chen 2019a).

Kartoittamistani monivaiheisissa hybridimalleista, jotka hyödyntävä datan hajottamisen tekniikoita, yksikään malleista ei käyttänyt WT:tä. Lisäksi mallit, jotka hyödynsivät WPD:tä käyttivät sitä ensimmäisenä hajottamismenetelmänä SDA:ssa. Tämä mahdollisesti viittaa siihen, että aalokepohjaiset hajottamisen menetelmät eivät yksinään takaa parhainta ennustuksen täsmällisyyttä.

EMD-pohjainen hajottaminen

Koska aalokepohjaiset menetelmät sisältävät puutteita, 1990-luvulla N. E. Huang ym. (1998) ehdottivat empiirisen moodin hajoamisen menetelmää (eng. Empirical Mode Decomposition, EMD) epälineaaristen ja ei-stationaaristen aikasarjojen analyysiin. EMD-menetelmässä aikasarja hajotetaan adaptiivisesti rajoitettuun määrään sisäisiä moodifunktioita (eng. Intrinsic Mode Functions, IMFs). Ne ovat luonteeltaan stationarisempia kuin alkuperäinen tuulen nopeuden aikasarja (Fu ym. 2020). EMD-menetelmällä on kaksi suurta ongelmaa: päätepiste-efekti ja moodien sekoittuminen. Päätepiste-efektissä (eng. endpoint effect) virheellisesti käytetään päätepisteenä arvoa, joka ei ole ääriarvo, sovitettaessa kuutiollista splinikäyrää (eng. cubic spline curve) (Liu ja Chen 2019a). Moodien sekoittumisessa (eng. mode mixing) signaalit, joilla on eri taajuus esiintyvät samassa IMF:ssä tai vastaavasti saman taajuuden omaavat signaalit esiintyvät eri IMF:ssä (Santhosh, Venkaiah ja Vinod Kumar 2018).

EMD-menetelmästä on tehty variaatioita, joilla on pyritty korjaamaan edellä mainittuja ongelmia. Ensemble-empiirisen moodin hajoamisessa, EEMD, lisätään valkoista kohinaa alkuperäiseen signaaliin. Tämä auttaa moodien sekoittumisen ongelmassa, mutta nyt ylimääräistä kohinaa esiintyy rekonstruoidussa signaalissa ja sen poistaminen on laskennallisesti raskasta (Qian ym. 2019). Tästä edelleen kehitetyssä täydennetyssä ensemble-empiirisen moodin hajoamisessa, CEEMD, lisätään myös valkoista kohinaa, mutta hyödyntäen vastakkaista polariteettia, joka seurauksena ylimääräisen kohinan poistaminen on tehokkaampaa (Yeh, Shieh ja Huang 2010). Torres ym. (2011) edelleen kehittivät CEEMDAN-menetelmän, joka pystyy vastaamaan edellä olevien menetelmien puutteisiin muun muassa vähentämällä laskennallista taakkaa ja kohinaa. Kuitenkin CEEMDAN-menetelmä kärsii jäännöskohinasta ja epäaidoista moodeista. Kehitetty ICEEMDAN-menetelmä pyrki ratkaisemaan nämä ongelmat (Song,

Wang ja Lu 2018; Colominas, Schlotthauer ja Torres 2014).

Kirjallisuudesta (Du ym. 2018) käy ilmi, että EEMD-menetelmä on parempi kuin EMD-menetelmä ja CEEMD-menetelmä toimii paremmin kuin kaksi edellä mainittua menetelmää. Qu ym. (2019) tuovat esille CEEMDAN-menetelmän vahvuudet suhteessa CEEMD-menetelmään. CEEMDAN-menetelmässä rekonstruoitu signaali on täydellinen ja vapaa kohinasta sekä se on laskennallisesti tehokkaampi. Tutkimuksessaan Song, Wang ja Lu (2018) vertasivat eri datan esikäsittelymenetelmiä. Tuloksista nähdään ICEEMDAN-menetelmä saavuttavan parempia tuloksia kuin CEEMDAN-menetelmän.

Kirjallisuudesta käy ilmi, että EMD-menetelmä ja sen variaatiot ovat käytetyimpiä datan hajottamisen tekniikoita hybridimalleissa (Ahmadi ja Khashei 2021). Kartoittamissani monivaiheisissa hybridimalleissa EMD-pohjaista hajottamista hyödyntävät mallit käyttivät usein CEEMD-, CEEMDAN- ja ICEEMDAN-menetelmiä. EMD- tai EEMD-menetelmää käytettäessä ne olivat osana SDA:ta joko ensimmäisenä tai toisena hajottamismenetelmänä.

Variaatiomoodin hajottaminen, VMD

EMD-menetelmällä ja sen variaatioilla on puutteensa, vaikka ne ovat yleisesti käytettyjä tuulen nopeuden ennustamisessa. Menetelmä on saanut kritiikkiä matemaattisen teoriataustan puutteesta, joten Dragomiretskiy ja Zosso (2014) esittelivät uuden hajottamismenetelmän variaatiomoodin hajottaminen (eng. Variational Mode Decomposition, VMD:n).

VMD:llä on vahva matemaattinen teoriatausta ja ei-rekursiivisena menetelmänä välttää virheet, jotka syntyvät rekursion aikana ja lopussa. VMD:sä ei-stationaarinen signaali hajotetaan joukoksi kaistarajoitettuja sisäisiä moodifunktioita (eng. Band-limited Intrinsic Mode Functions, BLIMFs) (Liu ja Chen 2019a). Kirjallisuudessa on näyttöä, että VMD tuottaisi täsmällisempiä ennustuksia kuin EEMD (D. Wang ym. 2017). VMD-menetelmä vaatii moodien sopivan lukumäärän asettamista etukäteen; liian pieni moodien lukumäärä puutteelliseen segmentointiin ja liian suuri määrä voi aiheuttaa kohinan kasvua tai moodien sekoittumista (Fu, Wang, Li ym. 2019). Optimaalinen-VMD, OVMD, jota ehdotti Zhang ym. (2017), pystyy itseadaptoituvana määrittämään moodien lukumäärän ja siten vastaamaan VMD:n puutteisiin.

Empiirinen aallokemuunnos, EWT

Toinen itsemukautuva datan hajottamisen menetelmä on empiirinen aallokemuunnos (eng. Empirical Wavelet Transform, EWT) (Gilles 2013). Liu, X.-w. Mi ja Y.-f. Li (2018) määrittelee EWT:n joukkona kaistapäästösuodattimia (eng. band pass filters), jotka on valittu signaalin spektristen ominaisuuksien perusteella ja jotka hajottavan alkuperäisen aikasarjan. EWT:n käyttö on lisääntynyt tuulen nopeuden aikasarjan hajottamisessa. Liu, X.-w. Mi ja Y.-f. Li (2018) tutkimuksessa EWT:n käyttö paransi ennustustuloksia verrattuna EMD- ja WPD-menetelmiin.

Hybridimallit, jotka hyödyntävät datan hajottamisen menetelmiä usein saavuttavat hyviä tuloksia, mutta useat näistä malleista voi olla laskennallisesti raskaita. Seuraavaksi esitellään datan esikäsittelymenetelmiä, joilla voi vähentää laskennallista taakkaa ilman mallin ennustuksen täsmällisyyden vähenemistä.

4.1.2 Ominaisuuksien valintaa ja ominaisuuksien irrottamista hyödyntävät mallit

Hybridimallit saattavat olla laskennallisesti raskaita. Eräs tapa vähentää laskennallista taakkaa on dimension pienentäminen. Tällöin datassa oleva tarpeeton ja toistuva tieto poistetaan, jonka seurauksena laskennallisuuden keventymisen lisäksi ennustuksen virhe pienenee (Liu ja Chen 2019a). Ominaisuuksien valinta ja ominaisuuksien irrottaminen ovat menetelmiä, joilla dimensiota voi pienentää.

Ominaisuuksien valinnassa valitaan mahdollisimman pieni osajoukko ominaisuuksia tiettyjen kriteerien mukaan alkuperäisestä syötteen ominaisuuksien joukosta (Liu ja Chen 2019a). Menetelmät voivat olla metaheuristisia optimaatiopohjaisia, korrelaatiopohjaisia, klusterointipohjaisia tai yhteiseen informaatioon pohjautuvia (Liu ja Chen 2019a). Lisäksi kirjallisuudesta mainitaan faasiavaruuden rekonstruktio menetelmä (eng. phase-space reconstruction, PSR) (Fu, Wang, Zhou ym. 2019) ja Gram-Schmidt ortogonaalimenetelmä, GSO, (C. Li ym. 2018).

Ominaisuuksien irrottamisessa taas kuvataan alkuperäinen ominaisuuksien joukko siten, että tuloksena on dimensioltaan pienempi joukko ominaisuuksia, jotka sisältävät mahdollisimman paljon alkuperäisen datan tietoa. Ominaisuuksien irrottamisessa voi syntyä uusia omi-

naisuuksia, piirteitä (eng. features) (Liu ja Chen 2019a). Tunnetuin lineaarinen menetelmä on pääkomponenttianalyysi (eng. principal component analysis, PCA) ja epälineaarista menetelmistä ytimen pääkomponenttianalyysi (eng. kernel principal component analysis, KPCA) (Liu ja Chen 2019a).

4.1.3 Kohinan poistaminen

Koska tuuli on hyvinkin satunnaista ja epävakaa, eräs tapa helpottaa ennustamista on etsiä ja vahvistaa tärkeää informaatiota datassa ja poistaa kohinaa. Yksi kohinan poistamisen menetelmistä liittyy vahvasti aiemmin esitettyyn hajottamisen menetelmiin 4.1.1. Tuulen nopeuden aikasarjan hajotettua useisiin osasarjoihin, komponentti tai komponentit, joilla suurin taajuus hylätään (Liu ja Chen 2019a). Eräs paljon kirjallisuudessa näkynyt kohinan poisto menetelmä on singulaarinen spektrianalyysi (eng. Singular Spectrum Analysis, SSA) (Liu, Mi ym. 2019; Xiao, Qian ja Shao 2017; Fu, Wang, Li ym. 2019). Menetelmässä yksi dimensionaalinen data muutetaan trajektorimatriisiksi (*Hankel matriisi*), josta valitaan merkittävimmät eli muutama ensimmäinen ominaisarvo rekonstruoimaan komponentit, joissa tärkeä informaatio on. Muuta pidetään kohinana. (Liu ja Chen 2019a).

4.1.4 Poikkeavien havaintojen havaitseminen

Jos data sisältää paljon epänormaalia informaatiota, joka ei kuvaa tuulen nopeuden aikasarjan todellista muutosta tai poikkeaa historiallisesta trendistä, niin nämä pisteet etsitään ja poistetaan. Tällaisilla poikkeavilla havainnoilla (eng. outliers) on virheellinen vaikutus ennustamiseen ja ennustus täsmällisyys jää pieneksi (Liu ja Chen 2019a). Ongelmana tuulen nopeuden aikasarjoissa on tunnistaa mikä on varsinaisesti poikkeava havainto ja mikä ei (Liu ja Chen 2019a). Ilmeisesti tämän vuoksi menetelmää ei juuri näkynyt tarkastelemassani kirjallisuudessa.

Yllä olevat datan käsittelytekniikat koskevat syötettä eli data käsitellään ennen ennustajaan syöttämistä. Poikkeavien havaintojen havaitsemista ja poistoa käytetään esikäsitelyn lisäksi myös jälkikäsitelyssä. Seuraavassa kappaleessa esitellään lisää datan jälkikäsitelyn tekniikoita.

4.2 Datan jälkikäsittelyn menetelmät hybridimalleissa

Datan jälkikäsittelyssä tarkoituksena on parantaa alustavaa ennustusta (Liu ja Chen 2019a). Näissä hybridimalleissa ensin suoritetaan ennustus ja lasketaan ennustusvirheet, analysoidaan ne ja luodaan niille ennustusmalli ja näiden pohjalta saadaan uudelleen tarkastettu malli (Ahmadi ja Khashei 2021). Tuulen nopeuden ennustamisessa jäännösvirheen mallintaminen (eng. residual error modeling) on Liu ja Chen (2019a) mukaan yleisimpiä jälkikäsittelyn tapoja. Esimerkiksi Y. Huang ym. (2019) käyttivät hybridimallissaan EEMD-LSTM-GRNN-BSO virheen korjauksessa GRNN- neuroverkkoa, joka oli optimoitu Brain Storm-optimointialgoritmillä. GRNN-BSO toteutettu virheen korjaus paransi ennustustuloksia.

Toinen kirjallisuudessa esiintyvä menetelmä jälkikäsittelyyn on suodatinpohjainen korjaus (eng. filter-based correction). Menetelmä perustuu oletukseen, että ennustetulla sarjalla on yhdenmukainen taajuuskaista alkuperäisen aikasarjan kanssa, joten alkuperäisen sarjan spektrin osaa käytetään suodattamaan virhe, odottamattomat ennustusarvot (Liu ja Chen 2019a). Erilaiset suodatinpohjaiset menetelmät näyttävät olevan toimiva ratkaisu monivaiheisten ennustusten tulosten parantamisessa suodattaen akkumuloituneesta ennustusvirheestä johtuvia ylimääräisiä komponentteja (Y. Li ym. 2019). Esimerkiksi Liu, Wu ja Li (2018) totesivat käänteisen EWT-menetelmän (eng. inverse EWT, IEWT) parantavan ennustuksen täsmällisyyttä ja stabiiliuutta. Y. Li ym. (2019) käyttivät WPF (eng. Wavelet packet filter)-menetelmää korjaamaan kertynyttä ennustusvirhettä. Suodatinpohjaisten korjausmenetelmät eivät ole kovin yleisiä hybridimalleissa (Liu ja Chen 2019a), vaikkakin niillä on huomattava vaikutus erityisesti monivaiheisten ennustusten stabiiliuuden ja täsmällisyyden parantamisessa (Liu, Wu ja Li 2018).

Seuraava kappale esittelee hybridimalleissa esiintyviä optimointialgoritmeja, joiden käyttö on yleistynyt viimeisen viiden vuoden aikana (Ahmadi ja Khashei 2021).

5 Optimointimenetelmät tuulen nopeuden ennustamisessa

Edelliset kappaleet ovat liittyneet tuulen nopeuden datan erilaisiin prosessointimenetelmiin. Optimointimenetelmät tuulen nopeuden ennustamisessa sen sijaan liittyvät vahvasti ennustajiin ja niiden erilaisten hyperparametrien, kuten oppimisnopeuden tai erän koon (eng. batch size) asettamiseen. Tämän lisäksi tuulen ennustamisessa optimointialgoritmeja hyödynnetään etsimällä sopivimmat painokertoimet eri ennustajille kootessa lopullista ennustustulosta yksittäisistä ennustuksista. Ennustukseen oleellisesti vaikuttaa ennustajan rakenne ja hyperparametrit, ja kirjallisuudessa on osoitettu, että erilaisilla optimointialgoritmeilla pystytään parantamaan ennustuksen täsmällisyyttä (Qu ym. 2019; Zhou, Sun ym. 2018).

5.1 Optimointi hybridimalleissa

Tuulen nopeuden ennustamisessa optimoinnissa hyödynnetään erilaisia metaheuristisia optimointialgoritmeja. Metaheuristiset algoritmit ovat yleiskäyttöisiä menetelmiä, joita voi hyödyntää laajassa joukossa erilaisia optimointiongelmia (Akbari ym. 2014). Ne ovat matemaattisia imitaatioita luonnonilmiöistä, jotka tehokkaasti löytävät optimaalisimman tuloksen välttämällä joutumista lokaaleihin ääriarvoihin (Mirjalili ja Lewis 2016). Eri algoritmeilla on erilaiset mekaniikat, mutta yleisellä tasolla niissä on samankaltaisuutta. Algoritmit alkavat ainakin yhdestä mahdollisesta ratkaisusta, tämän jälkeen ennalta määritetystä ratkaisuavaruudesta algoritmi etsii optimaalisia tuloksia ja iteratiivisesti päivittää vallitsevaa tilaa ja parametreja. Tätä toistetaan, kunnes algoritmin jokin lopetusehto täyttyy ja siten optimaalinen tulos on löytynyt (Liu, Chen ym. 2019). Vaikka optimointi algoritmit pyrkivät löytämään globaalin optimin, ne eivät välttämättä saavuta sitä.

Kirjallisuudessa nousi usein esille hyperparametrien optimointi tuulen nopeuden ennustamisessa. Alkuperäinen datajoukko jaetaan opetusjoukkoon, validointijoukkoon ja testijoukkoon. Parametrit, jotka saavuttavat pienimmän virheen validointijoukossa valitaan optimaaliseksi ja niitä käytetään lopullisessa ennustuksessa.

Fu, Wang, Li ym. (2019) käyttivät IHGWOSCA (eng. improved hybrid grey wolf optimizer-sine cosine algorithm) optimoimaan mallinsa PSR:n parametrit, ELM:n syötteen painot ja vakiot sekä piilotettujen neuroneiden lukumäärän. Tutkijoiden ehdotettu malli pystyi tehokkaasti optimoimaan valitut parametrit kolmen askeleen ennustuksessa. Prosentuaalinen absoluuttinen keskivirhe, MAPE, pieneni 91.34%, 86.92% ja 72.33% verrattuna malliin ilman IHGWOSCA optimointia. Memarzadeh ja Keynia (2020) optimoivat LSTM:n rakenteen ja syötteiden ominaisuuksien (eng. input feature) määrän Crow-etsintäalgoritmilla (eng. Crow Search Algorithm, CSA) ja vertasivat kyseistä algoritmia hiukkasparviontointiin (eng. Particle Swarm Optimization, PSO). Tuloksista on nähtävissä, että CSA:lla optimoitu mallit tuotti hieman parempia tuloksia kuin PSO:lla optimoitu malli. Tuulienergian ennustamisessa hyödynnettyjä metaheuristisia algoritmeja on useita. Liu, Chen ym. (2019) on tehnyt näistä laajan kartoituksen.

5.2 Monitavoiteoptimointi hybridimalleissa

Useat tutkijat toivat ilmi kirjallisuudessa, että tuulen ennustamisessa on tarve monitavoiteoptimoinnille, jotta voidaan saavuttaa sekä täsmällisiä että stabiileja tuloksia. Lisäksi monitavoiteoptimointi mahdollistaa myös mallein paremman keskinäisen vertailun (J. Wang ym. 2017). Monitavoiteoptimoinnissa yritetään löytää paras mahdollinen ratkaisu moneen eri, usein ristiriitaiseen, optimointiongelmaan (Liu, Chen ym. 2019). Monitavoiteoptimoinnin tuloksena saadaan yhden optimaalisen ratkaisun sijaan Pareto optimaalinen joukko eli joukko tuloksia, jotka eivät keskenään ole toista hallitsevimpia, mutta ovat parempi ratkaisujoukko kuin muut ratkaisuavaruuden ratkaisujoukot (Akbari ym. 2014).

Kuten aiemmin mainittiin, monitavoiteoptimoinnissa on yhden optimointifunktionsijaan useita. Tuulen ennustamisessa täsmällisempien ja stabiilimpien ennustusten saamiseksi usein yritetään minimoida keskineliövirhettä ja keskihajontaa samanaikaisesti tai ratkaista niin sanottu harha-varianssi dilemma (eng. bias-variance dilemma) eli yritetään sovittaa koko datajoukko samanaikaisesti vähentämällä sekä varianssia että datan perturbaation vaikutusta (Liu, Chen ym. 2019).

Monitavoiteoptimoinnilla saavutetaan parempia tuloksia usean askeleen ennustuksissa kuin yhdentavoitteen optimoinnilla. Esimerkiksi J. Wang ym. (2017) kehittivät valaiden optimointialgoritmista (eng. Whale Optimization Algorithm, WOA) ¹ monitavoiteoptimointialgoritmin MOWOA, joka saavutti parempia tuloksia kuin WOA ja aiemmin tutkimuksissa hyödynnetyt monitavoiteoptimoinnin algoritmit MOALO (eng. Multi-Objective Ant Lion Optimization) ja MODA (eng. Multi-Objective Dragonfly Algorithm).

Optimointimenetelmillä on positiivinen vaikutus tuulen nopeuden ennustamisen malleissa. Kuitenkin niiden varsinainen hyöty tulee esille malleissa, jossa hyödynnetään sekä datan esikäsittelyä. Esimerkiksi Du ym. (2017) tuulen nopeuden ennustustuloksien vertailussa (taulukko 2) on nähtävissä, että kaikissa datajoukoissa CEEMDAN-ENN saavutti huomattavasti parempi tuloksia kuin MOALO-ENN ja MOALO-ENN paransi vain hieman yksinkertaisen ENN-mallin tulosta Parhaimmat tulokset saavutti malli, jossa yhdistyi sekä datan esikäsittely että optimointialgoritmin hyödyntäminen.

Taulukko 2. Tuulen nopeuden ennustustulokset datajoukoille A-C (Du ym. 2017).

Datajoukko	Malli	RMSE(m/s)	MAPE(%)
A	ENN	0.8929	8.4867
	MOALO-ENN	0.8850	8.4280
	CEEMDN-ENN	0.5903	5.2033
	CEEMD-MOALO-ENN	0.4572	4.2121
B	ENN	0.8849	10.1538
	MOALO-ENN	0.8710	10.0826
	CEEMDN-ENN	0.5442	6.3156
	CEEMD-MOALO-ENN	0.4678	5.2429
C	ENN	0.9996	10.1368
	MOALO-ENN	0.9861	10.1082
	CEEMDN-ENN	0.6398	6.4831
	CEEMD-MOALO-ENN	0.5413	5.2145

Kuten datan käsittelymenetelmissä, niin myös optimointimenetelmissä tulisi tutkia ennus-

1. WOA algoritmi perustuu ryhävalaiden saalistuskäyttäytymisen analyysiin (Mirjalili ja Lewis 2016)

tajien parametrien optimoinnin itseadaptiivisuutta enemmän, jotta malleista saataisiin yleiskäyttöisempiä. Kirjallisuuskartoituksessani ei tullut vastaan itseadaptoituvia optimointialgoritmeja, joita olisi hyödynnetty tuulen nopeuden ennustamisessa.

6 Hybridimallit tuulen nopeuden monivaiheisessa ennustamisessa

Tässä kappaleessa esitellään tarkemmin muutamia hybridimalleja, joilla on saavutettu melko täsmällisiä ennustuksia monivaiheisissa ennustuksissa. Hybridimallit on valikoitu siten, että ne ovat lyhyen aikahorisontin malleja, niissä hyödynnetään sekä datan käsittelyä että optimointia ja monivaiheinen ennustus kattaa vähintään 5-askelta.

6.1 Mallien suorituskyvyn mittarit

Kirjallisuudessa mallien suorituskyvyn kvantitatiivisessa arvioinnissa usein käytetään kolmea tilastotieteellistä mittaria: absoluuttinen keskivirhe (eng. Mean Absolute Error, MAE), prosentuaalinen absoluuttinen keskivirhe (eng. Mean absolute percentage error, MAPE) ja keskineliövirheen neliöjuuri (eng. Root Mean Square Error, RMSE). Jatkossa tässä tutkielmassa viitataan näihin mittareihin niiden englanninkielisillä lyhenteillä MAE, MAPE ja RMSE. Taulukossa 3 on listattuna mittarit ja niiden kaavat.

Taulukko 3. Mittarit, joilla arvioidaan mallien suorituskykyä. (Niu ym. 2018)

Mittari	Selitys	Kaava
MAE	Absoluuttinen keskivirhe	$\frac{1}{N} \sum_{t=1}^N X(t) - X'(t) $
MAPE	Prosentuaalinen absoluuttinen keskivirhe	$\frac{1}{N} \sum_{t=1}^N \left \frac{X(t) - X'(t)}{X(t)} \right \times 100\%$
RMSE	Keskineliövirheen neliöjuuri	$\sqrt{\frac{1}{N} \sum_{t=1}^N [X(t) - X'(t)] ^2}$

Kaavoissa $X(t)$ kuvaa tuulen nopeuden aikasarjaa ja $X'(t)$ on ennustetut arvot.

6.2 Datan esikäsittelyä ja optimointia hyödyntävät hybridimallit

Qu ym. (2019) käyttävät mallissaan HDT-FPA-BPNN kaksiosaista hajotusmenetelmää, jota kutsuvat artikkelissaan hybridi hajottamismenetelmäksi (eng. Hybrid Decomposition Technique, HDT). Tässä alkuperäinen tuulen nopeuden aikasarja hajotetaan joukoksi sisäisiä moodifunktioita, IMF, joilla on eri frekvenssit, CEEMDAN-menetelmällä. Näistä funktiois-

ta IMF1 on kaikista ei-stationaarisin ja vaikein ennustettava, joten se hajotetaan edelleen EWT-menetelmällä saaden joukko eri moodeja, jotka ovat hieman vakaampia. Lisäksi he hyödyntävät kukkapölytysalgoritmia (eng. Flower Polination Algorithm, FPA) optimoimaan vastavirta-algoritmia käyttävän neuroverkon, BPNN, painot ja kynnykset (eng. weights and thresholds). FPA-BPNN syötteenä annetaan kaikki hajotetut komponentit, ja mallin lopullinen ennustustulos saadaan summaamalla nämä yksittäiset ennustukset. Tutkijoiden data oli kerätty Kiinassa Penglain ja Laizhoun kaupunkien tuulivoimaloista. Dataa kerättiin 15 minuutin välein 23. kesäkuuta 2013 20. heinäkuuta 2013 asti.

Tulokset vahvistavat jo tässä tutkielmassa aiemmin todettua: hybridimalli suoriutuu ennustamisesta huomattavasti paremmin kuin yksinkertaiset mallit. Esimerkiksi Laizhoun tuulivoimalasta kerätyssä datasta 5-astelehen monivaiheisessa ennustuksessa verrattaessa tutkijoiden ehdottamaa mallia yksinkertaiseen BPNN-malliin sekä hybridimalleihin FPA-BPNN, CEEMDAN-FPA.BPNN, EWT-FPA-BPNN MAPE arvoissa nähdään seuraavat pudotukset 68.70%, 61.03%, 31.66%,49.90%. Kaikissa vertailuissa ehdotettu malli pärjasi huomattavasti paremmin kuin muut mallit. Tulokset myös vahvistivat, että datan esikäsittelyllä on huomattavasti suurempi positiivinen vaikutus mallin ennustuksen täsmällisyyteen kuin esimerkiksi optimoinnilla. CEEMDAN ja EWT menetelmien väliltä ei voida sanoa selvästi kumpi on tehokkaampi.

Sun ym. (2018) ehdottamassa mallissa ASD-BSA-LRELM hyödynnetään adaptiivista sekundaarista hajottamismenetelmää (eng. Adaptive Secondary Decomposition, ASD). EEMD-menetelmällä hajotetaan alkuperäinen data joukoksi sisäisiä moodifunktioita, IMF, joista IMF1 hajotetaan edelleen käyttäen adaptiivista VMD-menetelmää, AVMD. Tutkijat optimoivat peruuttavan haun etsintäalgoritmeilla (eng. Backtracking Search Algorithm, BSA) äärimmäisen oppimiskoneen, LRELM, (eng. Leave-one-out-cross-validation Regularized Extreme Learning Machine, LRELM) parametrit syötteen painot ja piilotetut vakiot (eng. hidden biases). Datajoukot 1-4 oli kerätty Sisä-Mongoliasta 15 minuutin intervalleina viikon mittaisina jaksoina maaliskuu-, huhti- ja toukokuussa. Datajoukot A-D ovat Espanjasta Sotavento Galicia tuulivoimalasta 10 minuutin intervalleina ja jokaiselta vuodenajalta noin kahden viikon mittaisilta jaksoilta.

Tuloksista jälleen nähtävissä optimoinnin ja datan esikäsittelymenetelmien positiivinen vai-

kutus ennustus tuloksiin. Kaikille datajoukoille 6-, 9- ja 12- askeleen ennustuksille MAPE arvojen keskiarvoinen parantuminen verrattaessa ehdotettua ASD-BSA-LRELM mallia malleihin LRELM, AVMD-LRELM, ASD-LRELM on taulukossa 4.

Taulukko 4. MAPE arvojen keskimääräinen parantuminen (%) kaikille datajoukoille. (Sun ym. 2018)

Askel	LRELM	AVMD-LRELM	ASD-LRELM
6	62.63	33.11	27.38
9	59.38	34.98	24.95
12	52.74	27.56	20.38

Fu ym. (2020) hyödyntävät datan hajottamisessa EMD-menetelmää, jota on paranneltu aikaa varioivalla suodattimella (eng. Time Varying Filter, TVF). TVF:n tarkoituksena on vastata muun muassa EMD-menetelmässä esiintyvään moodien sekoittumisongelmaan (EMD-menetelmä kappaleessa 4.1.1). Alkuperäinen tuulen nopeuden sarja hajotetaan TVF-EMD-menetelmällä joukoiksi sisäisiä moodifunktioita, IMF. Tutkijat olivat todenneet, että usean datan käsittelymenetelmän yhdistäminen kasvattaa mallien monimutkaisuutta ja laskennallista taakkaa. He hyödyntävätkin sumean entropian teoriaa (eng. Fuzzy Entropy theory, FE theory), jotta voivat yhdistää TVF-EMD- menetelmän avulla hajotettuja komponentteja (IMF:t) ja saadaan näin pienemmän joukon komponentteja. SSA:n avulla tutkijat edelleen irrottivat joukoista oleellisia piirteitä.

Tutkijat myös yhdistivät kahta eri optimointialgoritmia, Harris Hawks-optimointialgoritmia (eng. Harris Hawks optimization, HHO) ja harmaasusien optimointialgoritmia (eng. Grey Wold Optimizer, GWO) tehden mutaatiopohjaisen optimointimenetelmän MHHOGWO:n. Ehdotettu optimointialgoritmi osoittikin konvergoituvan nopeammin ja globaalien optimaalisen arvon etsiminen parani. TVF-EMD-FE:n avulla saadut korkeamman frekvenssin komponentit syötettiin mallin PSR-KELM-osaan, joka oli optimoitu MHHOGWO-algoritmillä. Ja saatu matalan frekvenssin komponentti oli syötteenä ConvLSTM syväoppimisen neuroverkkoon. Dataa oli kerätty sekä 10-minuutin että tunnin intervalleissa. Keskityn tässä dataan, joka oli 10-minuutin välein kerättyä. Alueina oli Sotavento Galicia (SG) ja Beresford (BF).

Tuloksissa käy ilmi, että TVF-EMD yhdessä sumean entropiateorian kanssa kykenee vähentämään ennustettavien komponenttien lukumäärää ja siten vähentämään laskennallista taakkaa vaikuttamatta negatiivisesti lopulliseen ennustustulokseen. SSA-menetelmä ja MHHGWO vaikuttavat positiivisesti mallin toimintaan. Keskimääräinen MAPE arvon paraneminen 5-askeleen ennustuksessa Sotavento Galiciasta kerätylle datalle oli 47.95% ja Beresfordista kerätylle datalle oli 67.43%. Ja edelleen näiden keskiarvo on 57.92%.

Mielenkiintoista oli havaita, että malli saavutti parempia tuloksia Sotavento Galiciasta kerätylle datalle kuin Beresfordista, taulukko 5. Tarkastellessa datajoukkojen tilastollisia tietoja, kävi ilmi, että Sotavento Galiciasta kerätty data oli hieman tasaisempaa kun taas Beresfordin datan minimi ja maksimiarvojen välillä oli paljon suurempi väli, sekä keskihajonta oli suurempaa. Tämä voisi antaa osviittaa sille, että malli kykenee tunnistamaan vallitsevia trendejä hyvin, mutta jos data on hajanaisempaa, niin vaadittavien yksityiskohtien ennustaminen on heikompaa. Olisi mielenkiintoista, jos tutkijat olisivat tehneet 9-askeleen ja 12-askeleen ennustukset. Lisäksi tutkijat eivät olleet laittaneet esille CPU (s) aikaa, joka olisi ollut mielestäni mielenkiintoinen ottaen huomioon, että he olivat erityisesti pohtineet mallien laskennallista taakkaa ja siksi käyttäneet FE-teoriaa.

Taulukko 5. TVF-EMD-FE-SSA-MHHGWO-PSR-KELM-ConvLSTM mallin ennustustulosten MAPE(%) arvot. (Fu ym. 2020)

Datajoukko	1-askel	3-askel	5-askel
SG	2.2889	4.8798	5.0009
BF	3.0618	7.3745	7.9699

6.3 Datan esikäsittelyä, jälkikäsittelyä ja optimointia hyödyntävät hybridimallit

Kirjallisuudessa korostuvat datan esikäsittelytekniikoiden ja optimointimenetelmien hyödyntäminen. Datan jälkikäsittelyä, eli usein alustavan ennustuksen virheen korjaamista esiintyy hieman vähemmän. Tässä kappaleessa esittelen kaksi mallia, joissa on datan esikäsittelyn ja optimointimenetelmien lisäksi hyödynnetty datan jälkikäsittelyä.

Liu, Wu ja Li (2018) ehdottamassa hybridimallissa EWT-GWO-RELM-IEWT EWT hajottaa alkuperäisen tuulen nopeuden aikasarjan joukoksi stationaarisempia osasarjoja. Tutkijat hyödynsivät harmaa susi-optimoijaa RELM-mallin parametrien optimoinnissa. Lopullinen ennustettu sarja saadaan rekonstruoimalla osasarjojen yksittäiset ennustukset IEWT-algoritmin avulla. IEWT:n tarkoituksena on korjata mahdollisia odottamattomia mallin ennustamia arvoja. Tutkimuksen datana neljä tuulen aikasarjaa kiinalaisilta tuulivoimaloilta. Aikaintervallina on 10 minuuttia. Data on jaettu opetusjoukkoon, validaatiojoukkoon ja testausjoukkoon. Opetusjoukolla opetetaan RELM-malli, validointijoukkoa hyödynnetään etsiessä optimaaliset parametrit RELM-mallille ja testijoukon avulla arvioidaan mallin toimivuutta.

Tutkijoiden vertaillessa yksinkertaista RELM-mallia ja GWO-RELM-mallia. Optimointimenetelmä GWO ei kuurikaan parantanut tuloksia. 5- ja 7-askeleen MAPE arvot paranivat vain 1.35% ja 0.77%. Mielenkiintoista on, että verrattaessa ehdotettua hybridimallia EWT-GWO-RELM-IEWT samaan hybridimalliin ilman optimointia, EWT-RELM-IEWT, optimointimenetelmä huomattavasti paransi tuloksia. MAPE arvot paranivat 5- ja 7-askeleen ennustuksissa 69.98% ja 63.60%.

Tutkimuksen tarkoituksena oli myös selvittää onko IEWT-algoritmillä toteutetulla virheen korjauksella positiivisia vaikutuksia mallin toimivuuteen. Tuloksista nähdään verrattaessa ehdotettua hybridimallia samaan hybridimalliin ilman IEWT-algoritmin käyttöä, että IEWT-algoritmi parantaa ennustuksen täsmällisyyttä sekä stabiiliuutta. MAPE arvot paranivat 5- ja 7-askeleen ennustuksissa 28.43% ja 14.80%. Tutkijoiden ehdottama hybridimalli saavutti parhaimmat tulokset kaikista verrattavista malleista. Taulukossa 6 on nähtävissä mallin suorituskyvyn keskimääräinen arviointi kaikille datajoukoille.

Taulukko 6. EWT-GWO-RELM-IEWT-mallin suorituskyvyn arviointi. (Liu, Wu ja Li 2018)

Askel	MAPE (%)	MAE (m/s)	RMSE (m/s)
1	0.402525	0.032075	0.040025
5	1.591625	0.13005	0.1622
7	2.418775	0.1978	0.243075

Liu ja Chen (2019b) ehdottama hybridi mallissa käytetään resoluution muutosmenetelmiä dimension pienentämiseksi. Ensimmäinen näistä on keskiarvo-menetelmä ja toinen on syväoppimisen menetelmä, kaksikerroksinen pinottu harva autoenkoodari (eng. stacked sparse autoencoder, SSAE), jota käytetään ominaisuuksien irrottamisessa. Keskiarvomenetelmällä muutettu data hajotetaan hyödyntämällä EWT-algoritmia. Tutkijat käyttivät ennustamiseen kaksisuuntaista pitkäkestoista lyhytkestomuisti-neuroverkkoa (eng. Bidirectional Long Short-Term Memory network, BiLSTM). Ennustustulokset yhdistettiin hyödyntämällä MOMVO monitavoiteoptimointialgoritmia optimaalisten painojen löytämiseksi. Jäännösvirheen mallinnuksessa tutkijat käyttivät EWT-ORELM hybridimallia. ORELM (eng. outlier robust extreme learning machine) vähentää poikkeavien havaintojen vaikutusta ja siten pienentää ennustusvirhettä. Tutkimuksen data on Luoteis-Kiinasta, Xinjiangista. Dataa kerättiin kolmen sekunnin välein yhteensä 25 tunnin ajan 18.- 19. toukokuuta. Alkuperäinen data käsitti 30 000 otosta.

Tuloksista on nähtävissä, että jäännösvirheen mallintaminen EWT-ORELM hybridimallilla paransi huomattavasti ennustuksen täsmällisyyttä. Esimerkiksi 5-, 7- ja 10-askeleen ennustuksissa RMSE parani keskimäärin 83.2465%, 83.7204% ja 85.1175%. EWT-ORELM-malli saavutti myös muita jäännösvirheen mallintamiseen käytettyjä malleja parempia tuloksia. Tosin kaikilla verratuilla jäännösvirheen mallintamisen malleilla RELM, BPNN, ICEEMDAN-ARIMA, MOGWO-ELM oli positiivinen vaikutus tuloksiin. Ainoastaan SVM saavutti huonoja arvoja todennäköisen ylisovittumisen vuoksi.

Tutkimuksessa monitavoiteoptimointialgoritmia verrattiin yksitavoitteisiin optimointialgoritmeihin. Tuloksista on nähtävissä, että MOMVO saavutti pienimmän kumulatiivisen virheen. Tutkijoiden ehdottama hybridimalli saavuttaa erinomaisia tuloksia monivaiheisissa ennustuksissa. Tutkijat vertasivat malliaan vuosina 2018 ja 2019 julkaistuihin hybridimalleihin VDM-MOGWO-ELM, SSA-EMD-CNNSV, CEEMD-BPNN-BPNN_2-RBFNN-ENN-WNN-B-WCA ja WD-SampEn-VMD-MadaBoost-BFGS-WF. Tutkijoiden ehdottama malli saavutti huomattavasti parempia tuloksia, erityisesti monivaiheisen ennustuksen askelten kasvaessa. Taulukossa 7 on nähtävissä ehdotetun mallin suorituskyvyn arvioinnin keskiarvot 1-, 5- ja 10-askeleen ennustuksille.

Tämä tutkimus näyttää, että monitavoiteoptimointialgoritmien hyödyntäminen ja datan ha-

Taulukko 7. Hybridimallin suorituskyvyn keskimääräinen arviointi kaikille datajoukoille. (Liu ja Chen 2019b)

Askel	MAPE (%)	MAE (m/s)	RMSE (m/s)
1	0.402525	0.105633	0.164467
5	0.908867	0.153	0.2564
10	0.863267	0.151933	0.2618

jottamismenetelmien käyttö jäännösvirheen mallinnuksessa auttaa saavuttamaan hyviä tuloksia monivaiheisissa ennustuksissa.

Kuviossa 1 olevassa taulukossa on nähtävissä kaikkien yllä mainittujen hybridimallien suorituskyvyn arviointi. Taulukkoon on laskettu hybridimallien kaikkien datajoukkojen keskiarvot suorituskyvyn mittareille MAPE(%), MAE (m/s) ja RMSE (m/s). Taulukon tyhjä solu merkitsee, ettei kyseistä arvoa voitu laskea eli se ei ollut saatavilla tutkimuksesta. Vaikka näitä hybridimalleja ei voi suoraan verrata toisiinsa, niin kuitenkin on havaittavissa datan jälkikäsitteilyn suuri merkitys mallien suorituskyvyn parantumiselle. Hybridimallit EWT-GWO-RELM-IEWT ja EWT-BiLSTM-MOMVO-EWT-ORELM näyttävät saavuttavan täsmällisempiä ennustustuloksia.

Useaa kartoittamani hybridimallia, joita hyödynnettiin monivaiheisessa ennustamisessa, käytettiin vain 3-askeleen ennustuksissa tai 5-askeleen ennustuksissa. Käytännön sovelluksien kannalta olisi tärkeää, että monivaiheiset ennustukset pyrkisivät suurempiin askelmääriin, kuten 9-, 10- ja 12-askeleen ennustuksiin (Y. Li ym. 2019).

Malli	Aikaväli	1-askel				3-askel				5-askel			
		MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)	MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)	MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)
CEEMDAN-EWT-FPA-BP	15min	3,055	0,062	0,084	0,17	3,341	0,228	0,228	0,308	6,639	0,41		
ASD-BSA-LRELM	15min	4,615	0,191	0,3	0,42	7,25	0,673						
ASD-BSA-LRELM	10min	3,213	0,165	0,372	4		0,395						
TVF-EMD-FE-SSA-MHHOGWO-PSR-KELIM-ConvLSTM	10min	2,675	0,169	0,251	6,127	0,413	0,541	6,485	0,439	0,564			
EWT-GWO-RELM-IEWT	10min	0,403	0,032	0,04	0,983	0,079	0,1	1,592	0,13	0,162			
EWT-BILSTM-MOMVO-EWT-ORELM	3sek	0,588	0,106	0,164				0,856	0,153	0,256			

Malli	Aikaväli	6-askel				7-askel				9-askel			
		MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)	MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)	MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)
CEEMDAN-EWT-FPA-BP	15min												
ASD-BSA-LRELM	15min	11	0,665	0,955					15,25	0,898	1,198		
ASD-BSA-LRELM	10min	7,5	0,388	0,678					9	0,465	0,803		
TVF-EMD-FE-SSA-MHHOGWO-PSR-KELIM-ConvLSTM	10min												
EWT-GWO-RELM-IEWT	10min	1,978	0,162	0,201	2,419	0,198	0,243						
EWT-BILSTM-MOMVO-EWT-ORELM	3sek			0,259			0,258				0,267		

Malli	Aikaväli	10-askel				12-askel			
		MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)	MAPE (%)	MAE (m/s)	RMSE (m/s)	RMSE (m/s)
CEEMDAN-EWT-FPA-BP	15min								
ASD-BSA-LRELM	15min				21,25	1,215	1,565		
ASD-BSA-LRELM	10min				12,75	0,633	1,083		
TVF-EMD-FE-SSA-MHHOGWO-PSR-KELIM-ConvLSTM	10min								
EWT-GWO-RELM-IEWT	10min								
EWT-BILSTM-MOMVO-EWT-ORELM	3sek	0,863	0,152	0,262					

Kuvio 1. Koonti hybridimallien suorituskyvyn arvioinneista

7 Yhteenveto

Ilmastonmuutoksen hillitsemiseksi tärkeä toimenpide on energiatuotannon hiilidioksidipäästöjen vähentäminen lisäämällä uusiutuvia energialähteitä. Tuulienergian käyttö on nopeimmin kasvava uusiutuva energialähde. Kuitenkin tuulienergian lisääminen sähköverkkoon on haasteellista tuulen epävakaa ja epäsäännöllisen luonteen vuoksi. Jotta sähköverkon vakaus ja luotettavuus voidaan taata, tarvitaan tarkempia tuulivoiman ja tuulen nopeuden ennustuksia. Tuulen nopeuden ennustamista on tutkittu paljon viime vuosina. Ennustusmalleista hybridimallit ovat osoittautuneet suoriutuvan parhaiten tuulen nopeuden ennustamisesta. Hybridimalleissa koneoppimismenetelmään yhdistetään datan esikäsittelyä, jälkikäsittelyä ja/tai optimointimenetelmiä. Näitä menetelmiä on esitelty tässä tutkielmassa.

Erityisesti datan esikäsittelyllä, kuten hajottamisella ja kohinan poistolla, on merkittävä vaikutus ennustuksen täsmällisyyden parantumisella. Hybridimallien laskennallista taakkaa voidaan vähentää dimension pienentämisellä hyödyntämällä ominaisuuksien valintaa ja irrottamista. Jälkikäsittelymenetelmistä erityisesti jäännösvirheen mallintamisella on erinomainen vaikutus ennustajan suorituskykyyn. Tätä tulisi enemmän hyödyntää hybridimalleissa. Optimointimenetelmien ja monitavoiteoptimoinnin hyödyntäminen hybridimalleissa vaikuttaa positiivisesti mallin suorituskykyyn. Monitavoiteoptimoinnin avulla saavutetaan myös stabiilimpia ennustuksia.

Tutkielmassa tarkastelin monivaiheisia hybridimalleja ja menetelmiä, jotka parantavat näiden mallien suorituskykyä. Monivaiheinen ennustaminen on haasteellista, mutta tärkeää käytännön sovellusten kannalta, sillä se antaa energiajärjestelmille aikaa reagoida esimerkiksi poikkeustilanteissa. Kirjallisuuden perusteella on nähtävissä, että monivaiheiset hybridimallit, jotka hyödyntävät datan esikäsittelyä ja jälkikäsittelyä sekä optimointimenetelmiä saavuttavat parhaimpia tuloksia.

Tuulen nopeuden ennustamisessa käytetyt mallit ovat usein hyvin riippuvaisia kohteesta eivätkä yleiskäyttöisiä. Tulevaisuudessa olisi hyvä tutkia enemmän itsemukautuvia datan käsittelymenetelmiä ja optimointialgoritmeja. Lisäksi monivaiheisessa ennustamisessa tulisi pyrkiä suurempiin askelmääriin paljon tutkitun 3-asteleen sijaan. Myös siirtovaikutusoppimisen

käyttäminen tuulen nopeuden ennustamisessa voisi olla mielenkiintoinen tutkimusaihe.

Lähteet

Ahmadi, Mehrnaz, ja Mehdi Khashei. 2021. “Current status of hybrid structures in wind forecasting”. *Engineering Applications of Artificial Intelligence* 99:104133. ISSN: 0952-1976. <https://doi.org/https://doi.org/10.1016/j.engappai.2020.104133>.

Akbari, M., P. Asadi, M.K. Besharati Givi ja G. Khodabandehlouie. 2014. “13 - Artificial neural network and optimization”. Teoksessa *Advances in Friction-Stir Welding and Processing*, toimittanut Mohammad Kazem Besharati Givi ja Parviz Asadi, 543–599. Woodhead Publishing Series in Welding and Other Joining Technologies. Woodhead Publishing. ISBN: 978-0-85709-454-4. <https://doi.org/https://doi.org/10.1533/9780857094551.543>.

Cassola, Federico, ja Massimiliano Burlando. 2012. “Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output”. *Applied Energy* 99:154–166. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2012.03.054>.

Colominas, Marcelo A, Gastón Schlotthauer ja Maria E Torres. 2014. “Improved complete ensemble EMD: A suitable tool for biomedical signal processing”. *Biomedical Signal Processing and Control* 14:19–29.

Dhiman, Harsh S., Dipankar Deb ja Valentina Emilia Balas. 2020a. “Chapter 1 - Introduction”. Teoksessa *Supervised Machine Learning in Wind Forecasting and Ramp Event Prediction*, toimittanut Harsh S. Dhiman, Dipankar Deb ja Valentina Emilia Balas, 1–8. Wind Energy Engineering. Academic Press. ISBN: 978-0-12-821353-7. <https://doi.org/https://doi.org/10.1016/B978-0-12-821353-7.00012-0>.

———. 2020b. “Chapter 3 - Paradigms in wind forecasting”. Teoksessa *Supervised Machine Learning in Wind Forecasting and Ramp Event Prediction*, toimittanut Harsh S. Dhiman, Dipankar Deb ja Valentina Emilia Balas, 23–39. Wind Energy Engineering. Academic Press. ISBN: 978-0-12-821353-7. <https://doi.org/https://doi.org/10.1016/B978-0-12-821353-7.00014-4>.

Dragomiretskiy, K., ja D. Zosso. 2014. “Variational Mode Decomposition”. *IEEE Transactions on Signal Processing* 62 (3): 531–544. <https://doi.org/10.1109/TSP.2013.2288675>.

Du, Pei, Jianzhou Wang, Zhenhai Guo ja Wendong Yang. 2017. “Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting”. *Energy Conversion and Management* 150:90–107. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2017.07.065>.

Du, Pei, Jianzhou Wang, Wendong Yang ja Tong Niu. 2018. “Multi-step ahead forecasting in electrical power system using a hybrid forecasting system”. *Renewable Energy* 122:533–550. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2018.01.113>.

Fu, Wenlong, Kai Wang, Chaoshun Li ja Jiawen Tan. 2019. “Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM”. *Energy Conversion and Management* 187:356–377. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2019.02.086>.

Fu, Wenlong, Kai Wang, Jiawen Tan ja Kai Zhang. 2020. “A composite framework coupling multiple feature selection, compound prediction models and novel hybrid swarm optimizer-based synchronization optimization strategy for multi-step ahead short-term wind speed forecasting”. *Energy Conversion and Management* 205:112461. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2019.112461>.

Fu, Wenlong, Kai Wang, Jianzhong Zhou, Yanhe Xu, Jiawen Tan ja Tie Chen. 2019. “A Hybrid Approach for Multi-Step Wind Speed Forecasting Based on Multi-Scale Dominant Ingredient Chaotic Analysis, KELM and Synchronous Optimization Strategy”. *Sustainability* 11 (6). ISSN: 2071-1050. <https://doi.org/10.3390/su11061804>. <https://www.mdpi.com/2071-1050/11/6/1804>.

Gilles, Jerome. 2013. “Empirical wavelet transform”. *IEEE transactions on signal processing* 61 (16): 3999–4010.

Huang, Norden E, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung ja Henry H Liu. 1998. “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences* 454 (1971): 903–995.

Huang, Yuansheng, Lei Yang, Shijian Liu ja Guangli Wang. 2019. “Multi-Step Wind Speed Forecasting Based On Ensemble Empirical Mode Decomposition, Long Short Term Memory Network and Error Correction Strategy”. *Energies* 12 (10). ISSN: 1996-1073. <https://doi.org/10.3390/en12101822>.

Jiang, Ping, ja Zhenkun Liu. 2019. “Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting”. *Applied Soft Computing* 82:105587. ISSN: 1568-4946. <https://doi.org/https://doi.org/10.1016/j.asoc.2019.105587>.

Kavasseri, Rajesh G., ja Krithika Seetharaman. 2009. “Day-ahead wind speed forecasting using f-ARIMA models”. *Renewable Energy* 34 (5): 1388–1393. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2008.09.006>.

Kazienko, Przemyslaw, Edwin Lughofer ja Bogdan Trawinski. 2015. “Hybrid and ensemble techniques in soft computing: recent advances and emerging trends”. *Soft Comput* 19:3353–3355. ISSN: 1433-7479. <https://doi.org/https://doi.org/10.1016/j.enconman.2019.02.086>.

Li, Chaoshun, Zhengguang Xiao, Xin Xia, Wen Zou ja Chu Zhang. 2018. “A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting”. *Applied Energy* 215:131–144. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2018.01.094>.

Li, Yanfei, Huipeng Shi, Fengze Han, Zhu Duan ja Hui Liu. 2019. “Smart wind speed forecasting approach using various boosting algorithms, big multi-step forecasting strategy”. *Renewable Energy* 135:540–553. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2018.12.035>.

Liu, Hui, ja Chao Chen. 2019a. “Data processing strategies in wind energy forecasting models and applications: A comprehensive review”. *Applied Energy* 249:392–408. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2019.04.188>.

Liu, Hui, ja Chao Chen. 2019b. “Multi-objective data-ensemble wind speed forecasting model with stacked sparse autoencoder and adaptive decomposition-based error correction”. *Applied Energy* 254:113686. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2019.113686>.

Liu, Hui, Chao Chen, Xinwei Lv, Xing Wu ja Min Liu. 2019. “Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods”. *Energy Conversion and Management* 195:328–345. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2019.05.020>.

Liu, Hui, Ye Li, Zhu Duan ja Chao Chen. 2020. “A review on multi-objective optimization framework in wind energy forecasting techniques and applications”. *Energy Conversion and Management* 224:113324. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2020.113324>.

Liu, Hui, Xi-wei Mi ja Yan-fei Li. 2018. “Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network”. *Energy Conversion and Management* 156:498–514. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2017.11.053>.

Liu, Hui, Xiwei Mi ja Yanfei Li. 2018a. “An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm”. *Renewable Energy* 123:694–705. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2018.02.092>.

———. 2018b. “Comparison of two new intelligent wind speed forecasting approaches based on Wavelet Packet Decomposition, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Artificial Neural Networks”. *Energy Conversion and Management* 155:188–200. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2017.10.085>.

Liu, Hui, Xiwei Mi, Yanfei Li, Zhu Duan ja Yinan Xu. 2019. “Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression”. *Renewable Energy* 143:842–854. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2019.05.039>.

- Liu, Hui, Haiping Wu ja Yanfei Li. 2018. “Smart wind speed forecasting using EWT decomposition, GWO evolutionary optimization, RELM learning and IEWT reconstruction”. *Energy Conversion and Management* 161:266–283. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2018.02.006>.
- Liu, Mingshuai, Zheming Cao, Jing Zhang, Long Wang, Chao Huang ja Xiong Luo. 2020. “Short-term wind speed forecasting based on the Jaya-SVM model”. *International Journal of Electrical Power Energy Systems* 121:106056. ISSN: 0142-0615. <https://doi.org/https://doi.org/10.1016/j.ijepes.2020.106056>.
- Liu, Zhenkun, Ping Jiang, Lifang Zhang ja Xinsong Niu. 2020. “A combined forecasting model for time series: Application to short-term wind speed forecasting”. *Applied Energy* 259:114137. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2019.114137>.
- Marugán, Alberto Pliego, Fausto Pedro García Márquez, Jesus María Pinar Perez ja Diego Ruiz-Hernández. 2018. “A survey of artificial neural network in wind energy systems”. *Applied Energy* 228:1822–1836. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2018.07.084>.
- Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani ym. 2018. *Summary for Policymakers. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty*. <https://www.ipcc.ch/sr15/chapter/spm/>.
- Memarzadeh, Gholamreza, ja Farshid Keynia. 2020. “A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets”. *Energy Conversion and Management* 213:112824. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2020.112824>.
- Mirjalili, Seyedali, ja Andrew Lewis. 2016. “The Whale Optimization Algorithm”. *Advances in Engineering Software* 95:51–67. ISSN: 0965-9978. <https://doi.org/https://doi.org/10.1016/j.advengsoft.2016.01.008>.

Mosavi, Amir, Mohsen Salimi, Sina Faizollahzadeh Ardabili, Timon Rabczuk, Shahaboddin Shamshirband ja Annamaria R. Varkonyi-Koczy. 2019. “State of the Art of Machine Learning Models in Energy Systems, a Systematic Review”. *Energies* 12 (7). ISSN: 1996-1073. <https://doi.org/10.3390/en12071301>. <https://www.mdpi.com/1996-1073/12/7/1301>.

Nazir, Muhammad Shahzad, Fahad Alturise, Sami Alshmrany, Hafiz Nazir, Muhammad Bilal, Ahmad N Abdalla, P Sanjeevikumar, Ziad M Ali ym. 2020. “Wind generation forecasting methods and proliferation of artificial neural network: a review of five years research trend”. *Sustainability* 12 (9): 3778.

Niu, Tong, Jianzhou Wang, Kequan Zhang ja Pei Du. 2018. “Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy”. *Renewable Energy* 118:213–229. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2017.10.075>.

Niu, Xinsong, ja Jiyang Wang. 2019. “A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting”. *Applied Energy* 241:519–539. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2019.03.097>.

Qian, Zheng, Yan Pei, Hamidreza Zareipour ja Niya Chen. 2019. “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications”. *Applied energy* 235:939–953.

Qu, Zongxi, Wenqian Mao, Kequan Zhang, Wenyu Zhang ja Zhipeng Li. 2019. “Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network”. *Renewable Energy* 133:919–929. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2018.10.043>.

Santhosh, Madasthu, Chintham Venkaiah ja D.M. Vinod Kumar. 2018. “Ensemble empirical mode decomposition based adaptive wavelet neural network method for wind speed prediction”. *Energy Conversion and Management* 168:482–493. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2018.04.099>.

Secretariat, UNFCCC. 2016. *Report of the Conference of the Parties on its twenty-first session, held in Paris from 30 November to 13 December 2015. Addendum. Part two: Action taken by the Conference of the Parties at its twenty-first session*. <https://unfccc.int/documents/9097>.

Soman, S. S., H. Zareipour, O. Malik ja P. Mandal. 2010. “A review of wind power and wind speed forecasting methods with different time horizons”. Teoksessa *North American Power Symposium 2010*, 1–8. <https://doi.org/10.1109/NAPS.2010.5619586>.

Song, Jingjing, Jianzhou Wang ja Haiyan Lu. 2018. “A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting”. *Applied Energy* 215:643–658. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2018.02.070>.

Sun, Na, Jianzhong Zhou, Lu Chen, Benjun Jia, Muhammad Tayyab ja Tian Peng. 2018. “An adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine”. *Energy* 165:939–957. ISSN: 0360-5442. <https://doi.org/https://doi.org/10.1016/j.energy.2018.09.180>.

Torres, Maria E, Marcelo A Colominas, Gastón Schlotthauer ja Patrick Flandrin. 2011. “A complete ensemble empirical mode decomposition with adaptive noise”. Teoksessa *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 4144–4147. IEEE.

Valtioneuvosto. 2019. *Valtioneuvosto.fi*. Saatavilla WWW-muodossa, <https://valtioneuvosto.fi/marinin-hallitus/hallitusohjelma/hiilineutraali-ja-luonnon-monimuotoisuuden-turvaava-suomi>, viitattu 10.02.2021.

Wang, Deyun, Hongyuan Luo, Olivier Grunder ja Yanbing Lin. 2017. “Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction”. *Renewable Energy* 113:1345–1358. ISSN: 0960-1481. <https://doi.org/https://doi.org/10.1016/j.renene.2017.06.095>.

Wang, Jianzhou, Pei Du, Tong Niu ja Wendong Yang. 2017. “A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting”. *Applied Energy* 208:344–360. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2017.10.031>.

Wang, Jianzhou, ja Jianming Hu. 2015. “A robust combination approach for short-term wind speed forecasting and analysis – Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model”. *Energy* 93:41–56. ISSN: 0360-5442. <https://doi.org/https://doi.org/10.1016/j.energy.2015.08.045>.

Wang, Jianzhou, Yiliao Song, Feng Liu ja Ru Hou. 2016. “Analysis and application of forecasting models in wind power integration: A review of multi-step-ahead wind speed forecasting models”. *Renewable and Sustainable Energy Reviews* 60:960–981. ISSN: 1364-0321. <https://doi.org/https://doi.org/10.1016/j.rser.2016.01.114>.

Wang, Jujie, ja Yaning Li. 2018. “Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy”. *Applied Energy* 230:429–443. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2018.08.114>.

Xiao, Liye, Feng Qian ja Wei Shao. 2017. “Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm”. *Energy Conversion and Management* 143:410–430. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2017.04.012>.

Yang, Wendong, Jianzhou Wang, Haiyan Lu, Tong Niu ja Pei Du. 2019. “Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: A case study in China”. *Journal of Cleaner Production* 222:942–959. ISSN: 0959-6526. <https://doi.org/https://doi.org/10.1016/j.jclepro.2019.03.036>.

Yeh, Jia-Rong, Jiann-Shing Shieh ja Norden E Huang. 2010. “Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method”. *Advances in adaptive data analysis* 2 (02): 135–156.

Zhang, Chu, Jianzhong Zhou, Chaoshun Li, Wenlong Fu ja Tian Peng. 2017. “A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting”. *Energy Conversion and Management* 143:360–376.

Zhao, Jing, Zhen-Hai Guo, Zhong-Yue Su, Zhi-Yuan Zhao, Xia Xiao ja Feng Liu. 2016. “An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed”. *Applied Energy* 162:808–826. ISSN: 0306-2619. <https://doi.org/https://doi.org/10.1016/j.apenergy.2015.10.145>.

Zhou, Jianzhong, Han Liu, Yanhe Xu ja Wei Jiang. 2018. “A Hybrid Framework for Short Term Multi-Step Wind Speed Forecasting Based on Variational Model Decomposition and Convolutional Neural Network”. *Energies* 11 (9). ISSN: 1996-1073. <https://doi.org/10.3390/en11092292>. <https://www.mdpi.com/1996-1073/11/9/2292>.

Zhou, Jianzhong, Na Sun, Benjun Jia ja Tian Peng. 2018. “A Novel Decomposition-Optimization Model for Short-Term Wind Speed Forecasting”. *Energies* 11 (7). ISSN: 1996-1073. <https://doi.org/10.3390/en11071752>.

Zhou, Junyi, Jing Shi ja Gong Li. 2011. “Fine tuning support vector machines for short-term wind speed forecasting”. *Energy Conversion and Management* 52 (4): 1990–1998. ISSN: 0196-8904. <https://doi.org/https://doi.org/10.1016/j.enconman.2010.11.007>.