

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA ELEKTROTECHNICKÁ
KATEDRA ELEKTROMAGNETICKÉHO POLE



**ODHAD PROFILU REFRAKTIVITY POMOCÍ
UMĚLÝCH NEURONOVÝCH SÍTÍ**

DISERTAČNÍ PRÁCE
ING. MARTIN MUDROCH

DOKTORSKÝ STUDIJNÍ PROGRAM: ELEKTROTECHNIKA A INFORMATIKA

STUDIJNÍ OBOR: RADIOELEKTRONIKA

ŠKOLITEL: PROF. ING. MILOŠ MAZÁNEK, CSc.

ŠKOLITEL SPECIALISTA: PROF. ING. PAVEL PECHAČ, Ph.D.

PRAHA, ČERVEN 2014

Obsah

Seznam obrázků	iv
Seznam tabulek	vi
Seznam zkratk	viii
1 Úvod	1
2 Metody predikce, klasifikace a analýzy dat	3
2.1 Analýza dat	3
2.1.1 Statistické metody	3
2.1.1.1 Shluková analýza	3
2.1.1.2 Support Vector Machines	4
2.1.2 Neuronové sítě pro analýzu dat	4
2.2 Klasifikace	5
2.2.1 Rozhodovací stromy	5
2.2.2 Fuzzy logika	5
2.2.3 Neuronové sítě pro klasifikaci	5
2.3 Predikce	5
2.3.1 Nelineární regrese a predikce časových řad	6
2.3.2 Nelineární data-fitting	6
2.3.3 Neuronové sítě pro predikci	6
2.4 Výběr vhodných metod	7
3 Současný stav problematiky	8
3.1 Oblast radiových spojů	8
3.1.1 Publikované metody pro odhad refraktivity v troposféře	8
3.2 Šíření radiového signálu v troposféře	9
3.2.1 Meteorologické jevy ovlivňující šíření	9
3.2.1.1 Srážky	9
3.2.1.2 Přízemní oblačnost	10
3.2.1.3 Mlha	10
3.2.1.4 Ostatní jevy	10

3.2.2	Útlum způsobený průchodem elektromagnetické vlny prostředím	11
3.2.3	Útlum hydrometeory	11
3.2.4	Atmosférické turbulence	11
3.2.5	Troposférická refrakce	12
3.3	Metody výpočtu vlivu troposférické refrakce na radiové spoje . . .	14
3.4	Oblast bezdrátových optických spojů	15
3.5	Vlivy prostředí na FSO spoje	15
3.5.1	Extinkce	16
3.5.2	Rozptyl	17
3.5.2.1	Rayleighův rozptyl	17
3.5.2.2	Mieův rozptyl	18
3.5.3	Turbulence	19
3.6	Zhodnocení současného stavu	19
4	Cíle práce	20
5	Neuronové sítě a jejich použití	22
5.1	Historie	22
5.2	Druhy neuronových sítí	25
5.2.1	Perceptron	26
5.2.1.1	Aktivační funkce neuronu	26
5.2.1.2	Aplikační možnosti jednovrstvé sítě perceptronovského typu	28
5.2.1.3	Učení jednovrstvé sítě perceptronovského typu . .	29
5.2.2	Multi-layer Perceptron	31
5.2.2.1	Učení MLP sítě	32
5.2.2.2	Algoritmus back-propagation (BP)	33
5.2.3	Kohonenova síť - SOM	36
5.2.3.1	Učení Kohonenovy sítě	38
5.2.3.2	Vybavování Kohonenovy sítě	40
5.2.4	Rekurentní síť	40
5.3	Aplikace neuronových sítí	41
6	Odhad refraktivity pro pozemní radiové spoje	43
6.1	Dílčí cíle	43
6.2	Popis měřicího spoje	43
6.3	Rozbor měřených dat	45
6.4	Klasifikace množství srážek pomocí radiových spojů	49
6.4.1	Vstupní data	49
6.4.2	Učící proces	50
6.4.3	Vyhodnocení učícího procesu	51
6.4.4	Dílčí závěry	52

6.5	Konstrukce učících množin pro meteorologické aplikace	55
6.6	Odhad refraktivity a jejího výškového profilu	56
6.6.1	Odhad profilu refraktivity	57
6.6.2	Tvorba učících množin	58
6.7	Vyhodnocení kvality odhadu profilu refraktivity	59
6.8	Shrnutí použité metodiky	60
6.9	Dílčí závěr	61
7	Návrh metodiky predikce vlivu počasí na FSO spoje	63
7.1	Dílčí cíle	63
7.2	Popis experimentu	63
7.3	Analýza vstupních dat	64
7.3.1	Shluková analýza pomocí SOM sítě	65
7.4	Návrh metodiky modelování atmosferických vlivů na FSO spoj . .	67
7.5	Popis jednotlivých kroků metodiky	68
7.5.1	Příprava vstupních dat	68
7.5.2	Výběr architektury a velikosti neuronové sítě	69
7.5.3	Konstrukce učících množin	72
7.5.4	Učení neuronové sítě	72
7.5.5	Vyhodnocení jednotlivých neuronových sítí	72
7.6	Dílčí závěr	76
7.7	Doporučení pro další zlepšení	76
8	Závěr	78
	Literatura	79

Seznam obrázků

3.1	Standardní refrakce	13
3.2	Subrefrakce	13
3.3	Superrefrakce	14
3.4	Přízemní vlnovodný kanál	14
3.5	Vyvýšený vlnovodný kanál	14
5.1	Biologický neuron, který se stal předlohou pro umělý neuron	23
5.2	Grafická reprezentace klasifikace funkce AND a problému s funkcí XOR	25
5.3	Schéma funkce jednoho perceptronu	26
5.4	Průběh funkce unipolární sigmoidy podle (5.8) s parametrem sigmoidy $\alpha = 1$	27
5.5	Průběh funkce bipolárních sigmoid: $f1 = \tanh(x)$, $f2 =$ funkce (5.10)	27
5.6	Srovnání funkcí $\tansig()$ a $\logsig()$	28
5.7	Heavisideova funkce	28
5.8	Vícevrstvá síť perceptronovského typu (Multi-Layer Perceptron - MLP)	32
5.9	Architektura Kohonenovy sítě - SOM	37
5.10	Příklad možných definic uspořádání vzdáleností ve výstupní vrstvě SOM: a) čtvercové okolí, b) šestiúhelníkové okolí	38
5.11	Průběh funkce mexického klobouku - rovnice (5.31)	38
5.12	Průběh schodové funkce ostrého okolí pro $p = 3$	38
5.13	Nerrandova kanonická forma	41
6.1	Terénní profil experimentálního spoje s naznačenými jednotlivými spoji	44
6.2	Schéma uspořádání měřicích zařízení na stožáru Poděbrady	46
6.3	Stožár Poděbrady s přijímači a meteosenzory - spodní pohled ze země	47
6.4	Stožár Poděbrady - skříň se zařízením pro načítání dat a 1. čidlo teploty a vlhkosti	47
6.5	Průběh útlumů měřicích spojů za deště	47
6.6	Průběh radiových útlumů jednoho dne - 1/12/2007	48
6.7	Průběh útlumů jednoho dne - 10/12/2007	48

6.8	Průběh útlumů jednoho dne - 23/12/2007	49
6.9	Vyhodnocení MSE pro neuronové sítě se dvěma skrytými vrstvami a testovací množinu MIX	53
6.10	Vyhodnocení MSE pro neuronové sítě se dvěma skrytými vrstvami a testovací množinu DAY	53
6.11	Vyhodnocení MSE pro neuronové sítě se třemi skrytými vrstvami a testovací množinu MIX	54
6.12	Vyhodnocení MSE pro neuronové sítě se třemi skrytými vrstvami a testovací množinu DAY	54
6.13	Příklad rozptylu dat pro nejlepší síť a trénovací množinu - 100x50 a množinu MIX2. Proloženo přímkou, která znázorňuje skutečný průběh (modrá) a červenou, která znázorňuje ideální závislost. . .	56
6.14	Histogram chyby pro odhad gradientu refraktivity pro síť 30x15 .	57
6.15	Průběh funkcí (6.3) a (6.3) použitých k normalizaci	59
7.1	Histogram naměřeného útlumu	65
7.2	Histogram naměřené teploty	65
7.3	Histogram naměřené vlhkosti	65
7.4	Časový průběhu srážek	65
7.5	Zobrazení míry excitace jednotlivých neuronů	66
7.6	Mapy vah pro všechny vstupy	67
7.7	Ukázka histogramu teploty před provedením normalizace	69
7.8	Histogram teploty po normalizaci funkcí (7.2)	69
7.9	Histogram chyby pro síť 25x15	70
7.10	Architektura neuronové sítě se zpožděnými daty útlumů	71
7.11	Architektura neuronové sítě se zpožděnými meteorologickými daty	71
7.12	Časový průběh velikosti chyby odhadu pro neuronovou síť 14x9 v konfiguraci C6	74
7.13	Průměrná rychlost větru v závislosti na teplotě pro simulační chyby větší než 1 dB.	76

Seznam tabulek

6.1	Nejlepší MSE dosažené pro všechny sítě a učící množiny	52
6.2	Zastoupení jednotlivých tříd srážek v učících množinách	55
6.3	Kvalita klasifikace pro všechny učící množiny a neuronové sítě . . .	56
6.4	Vyhodnocení přesnosti odhadu gradientu refraktivity	57
6.5	Vyhodnocení odhad celkového gradientu refraktivity pro všechny trénovací množiny	59
6.6	Vyhodnocení odhad jednotlivých pater gradientu refraktivity pro učící množinu TS1	60
6.7	Vyhodnocení odhad jednotlivých pater gradientu refraktivity pro učící množinu TS2	60
6.8	Vyhodnocení celkového odhadu gradientu refraktivity pro obě tré- novací množiny	60
7.1	Vyhodnocení prvního odhadu rozměru neuronové sítě pro model FSO	70
7.2	Vyhodnocení chyby pro síť o velikosti 14x9 pro odhad chování spoje WaveBridge	74
7.3	Vyhodnocení chyby pro síť o velikosti 17x11 pro odhad chování spoje WaveBridge	74
7.4	Vyhodnocení chyby pro síť o velikosti 9x6 pro odhad chování spoje WaveBridge	75
7.5	Vyhodnocení chyby pro nejlépe natrénovanou síť 17x11 pro spoj FlightStrata	75

Seznam zkratek

ADALINE	ADaptive LInear NEuron
ANN	Artificial Neural Network
BP	Back-Propagation
FF-ANN	Feed-Forward Artificial Neural Network
FSO	Free Space Optics
GDR	Generalized Delta-Rule
MADALINE	Many ADALINE
MCP	McCulloch-Pittsův neuron
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
NARX	Nonlinear Autoregressive Network with Exogeneous Inputs
RSSI	Received Signal Strength Indicator
SOM	Self Organizing Maps
SVM	Support Vector Machines
TD-NN	Time-Delay Neural Network
UI	Umělá inteligence

Kapitola 1

Úvod

Tato práce se zaměřuje na analýzu možností zpracování naměřených meteorologických dat v kombinaci s informacemi o stavu radiových (RF) a optických (FSO) bezdrátových spojů. Na datech získaných ze specifických experimentů je navrhována metodika pro aplikaci vybraných přístupů známých z oblasti data miningu resp. umělé inteligence pro zmíněné oblasti bezdrátové komunikace. Experimenty nejsou původně cíleny na zpracování pomocí umělé inteligence (UI), ale spíše na ověřování a zpřesnění existujících fyzikálních modelů definovaných doporučeními ITU-R. Použití UI je tedy novým přístupem k problematice.

V práci jsou definovány postupy pro aplikaci neuronových sítí, které byly mezi ostatními metodami zvoleny jako nejvhodnější pro řešení dále popsaných problémů. Následně jsou shrnuta doporučení pro aplikaci těchto sítí pro dané třídy problémů. Řešená problematika je dobře převoditelná do oblastí predikce a klasifikace. Pro řešení obecných problémů v těchto oblastech jsou často úspěšně využívány neuronové sítě. Výhodou těchto sítí je to, že není nutná podrobná znalost konkrétních fyzikálních závislostí dat získaných v experimentech.

První část práce obsahuje přehled použitelných standardních resp. statistických metod jako i dalších, potenciálně vhodných, metod umělé inteligence, které mohou být použity pro problémy klasifikace a predikce. Jedná se o zpracování experimentálních dat, která mohou být zatížena jak chybou měření, tak i meteorologickými vlivy, které experiment není schopen vhodným způsobem zachytit. V kapitole jsou na základě popsaných vlastností jako nejvhodnější prostředek pro řešení zadané problematiky zvoleny umělé neuronové sítě a tento výběr je podrobněji zdůvodněn.

Druhá část práce obsahuje shrnutí současného stavu problematiky s popisem jednotlivých atmosferických vlivů, které jsou v této práci uvažovány při tvorbě metodiky pro radiové i bezdrátové optické spoje.

Třetí část práce obsahuje podrobný rozbor postupů a architektur neuronových sítí, které jsou obecně použitelné pro nelineární klasifikaci a predikci nejrůznějších složitých závislostí. Jsou zde popsány i Kohonenovy samoorganizující mapy (SOM), které jsou používány pro analýzu vstupních dat jako doplněk statistických

metod. Základním stavebním prvkem aplikací neuronových sítí ve zmíněné problematice jsou dopředné neuronové sítě (feed-forward neural networks). Zároveň jsou zde diskutovány vlastnosti rekurentních neuronových sítí, které mají dobré využití například v oblasti predikce časových řad a v obdobných závislostech, ke kterým se řešení modelu přibližuje.

Čtvrtá část práce obsahuje popis navržené metodiky aplikace neuronových sítí na data získaná z experimentálního radiového spoje (o délce přibližně 50 km, na frekvenci 10 GHz). Jedná se o unikátní experimentální spoj s jedním vysílačem a více přijímači v různých výškách nad terénem na jednom stožáru. Tento stožár je rovněž doplněn o meteorologická čidla. Kapitola popisuje fyzikální jevy a meteorologické pochody, které ovlivňují šíření radiového signálu v nižších vrstvách troposféry. Vzhledem k tomu, že spoje procházejí různými vrstvami atmosféry, ve kterých se mohou lišit podmínky pro šíření, byla data z tohoto experimentu použita k odhadu vertikálního profilu refraktivity. Získaná odhadnutá hodnota refraktivity slouží jako pomůcka k identifikaci jednotlivých jevů z hlediska šíření radiového signálu v atmosféře jako jsou například subrefrakce, superrefrakce a vlnodný kanál. Dále je zde diskutována aplikace neuronové sítě na odhad intenzity dešťových srážek na základě kombinace naměřených útlumů.

Pátá část je zaměřena na rozšíření a zpřesnění metodiky užití neuronových sítí tentokrát na kratší bezdrátový optický spoj (FSO). Neuronové sítě jsou v této oblasti využity pro predikci útlumu těchto komunikačních spojů. Tyto sítě jsou zde aplikovány na dva nezávislé optické spoje (s různými možnostmi měření útlumových dat) umístěné v městském prostředí o délce do 500 m. Vstupními daty neuronové sítě jsou meteorologické údaje a útlum FSO spojů. Zpětně jsou vyhodnoceny požadavky na spoj, které musí být splněny, aby bylo možné pomocí neuronové sítě predikovat chování FSO spoje. V této kapitole jsou dále diskutovány rozdíly v přesnosti predikce pro různé architektury neuronových sítí a je zde popsán vývoj modelu směrem k rekurentním neuronovým sítím.

Na závěr je provedeno shrnutí a diskuse dosažených cílů jak pro oblast odhadu refraktivity, tak i pro neuronové modelování vlivu počasí na FSO spoje. Jsou zde identifikovány problémy, které ovlivňují přesnost uvedených neuronových metod a jsou zmíněny i možnosti, které mohou do určité míry pomoci tyto problémy překonat.

Kapitola 2

Metody predikce, klasifikace a analýzy dat

Tato kapitola obsahuje srovnání statistických metod a přístupů umělé inteligence, které jsou známé a lze je obecně použít pro řešení problémů klasifikace, predikce a analýzy dat. Tyto obecné operace jsou základem pro řešení problémů popsanych v dalších kapitolách.

2.1 Analýza dat

Analýza dat vždy předchází zpracování naměřených experimentálních hodnot a jejím cílem je získat obecné informace o jejich struktuře, jejich rozsazích a případných vazbách, které nejsou předem zřejmé. Typickým a nejjednodušším prostředkem jsou statistické metody. Existují i oblasti umělé inteligence, které jsou schopny analyzovat velké soubory dat s málo zřejmými vnitřními vazbami. Typickým představitelem je například strojové učení, respektive jeho aplikace v neuronových sítích.

2.1.1 Statistické metody

Typickým úkolem statistického zpracování, je zjištění druhu statistického rozložení měřených fyzikálních veličin, jejich střední hodnoty, rozptylu a i dynamiky v čase. Tato znalost je pak využitelná při další práci s naměřenými hodnotami a je nezastupitelná jinými prostředky. Existují i pokročilé statistické metody, kterými lze proniknout hlouběji do vnitřních vztahů uvnitř dat. Takovými metodami může být například vícerozměrná shluková analýza, Support Vector Machines (SVM) a další.

2.1.1.1 Shluková analýza

Cílem shlukové analýzy je nalezení různých množin vstupních vektorů naměřených dat, jejichž prvky si jsou do určité míry podobné. Bylo by tedy možné

nalézt typické zástupce každé množiny a tím snížit mocnost naměřených dat vynecháním některých prvků z nalezených skupin. Nalezený shluky rovněž mohou představovat nějakou těžko zřejmou fyzikální závislost v datech. Existuje celá řada algoritmů, které lze použít na vyhledávání shluků ve vícerozměrných datech. Problémem statistických metod je nalezení statistických hlavních komponent, pomocí kterých lze data transformovat, aby došlo k co největšímu snížení počtu dimenzí experimentálních dat, protože s počtem dimenzí roste i výpočetní náročnost dalších výpočtů. Přehled algoritmů shlukové analýzy lze najít například v [1] - jsou zde shrnuty jak statistické postupy tak i algoritmy využívané v neuronových sítích a strojovém učení.

2.1.1.2 Support Vector Machines

Princip strojového učení u SVM je založen na využití jádrových algoritmů a dokáže hledat separující funkce nad vícerozměrnými daty. Vyhledávání separující funkce probíhá pomocí podpůrných vektorů. Pokud jsou množiny lineárně neseparovatelné v jednom prostoru je použita vhodná jádrová transformace a je zvýšena dimenze separující funkce. Tento přístup dokáže velmi dobře hledat separující funkce, ale jeho problémem je relativně velká výpočetní náročnost. Popis algoritmu shlukové analýzy je uveden například v [2] nebo [3]. Výhody tohoto přístupu oproti ostatním algoritmům se projeví až při relativně velkém množství rozlišovaných atributů a pro případ této práce je přístup pomocí SVM zbytečně náročný.

2.1.2 Neuronové sítě pro analýzu dat

Neuronové sítě nelze použít jako základní prostředek pro analýzu dat, protože jejich aplikace ve většině případů už vyžaduje určitou základní přípravu dat. Jejich aplikace je ale v této oblasti přínosná například pro shlukovou analýzu. Kohonenovy samoorganizující mapy jsou typickým prostředkem, který dokáže analyzovat vnitřní vztahy uvnitř naměřených souborů dat (přehled postupů lze najít například v [4]). Tyto sítě jsou rovněž dobrým prostředkem, kterým může být analyzována závislost jednotlivých měřených veličin - obdobně jako statistickou korelací. Pro Kohonenovy mapy bylo vytvořeno mnoho vizualizačních metod, které pomáhají vyhodnocovat nalezené shluky v datech. U tohoto typu sítí není velkým problémem dimenze vstupních dat, jako například u statistických metod. Náročnost výpočtu shluků pomocí SOM je ve větší míře závislá spíše na kvalitě (ostrosti hran) jednotlivých shluků a samozřejmě i na návrhu metriky použité v SOM síti. Bližší informace o obecných postupech v sítích SOM je uveden v samostatné kapitole věnované neuronovým sítím.

2.2 Klasifikace

Klasifikační problémy řeší problémy jak rozdělit vícerozměrná data do skupin, která mají nějakou společnou vlastnost.

2.2.1 Rozhodovací stromy

V případech, kdy jsou závislosti popsitelné pomocí určitých pravidel, lze na klasifikaci těchto závislostí použít rozhodovací stromy. Pro aplikaci rozhodovacích stromů v zadané problematice je ale velmi složité nalézt pro množinu vstupních dat konkrétní pravidla, která by byla obecně platná. Problémem jsou například případy, kdy pro stejná meteorologická data poskytují experimenty rozdílné výsledky vlivem nemožnosti podchytit všechny atmosferické vlivy naměřenými meteodaty. V takových problémech by konstrukce rozhodovacích stromů selhala. Přehled přístupů k tvorbě rozhodovacích stromů lze najít například v [5].

2.2.2 Fuzzy logika

Fuzzy logiku lze považovat za rozšíření standardní logiky o míru pravděpodobnosti. Tento přístup opět naráží na problém s definicí příslušných pravidel stejně jako v případě rozhodovacích stromů. Nicméně na základě publikace [6] lze předpokládat, že pomocí fuzzy logiky lze zmírnit některé problémy rozhodovacích stromů (především právě neurčitost pro sporná vstupní data).

2.2.3 Neuronové sítě pro klasifikaci

Neuronové sítě s vhodně navrženou architekturou mohou sloužit jako univerzální aproximátory a klasifikátory obecných nelineárních vícerozměrných funkcí. To bylo v literatuře [7] již dokázáno. Pro svoji funkčnost nepotřebují žádné předchozí znalosti fyzikálních zákonitostí, které se v používaných datech vyskytují. Jejich problémem je ale jejich specifická pro konkrétní problém. Neuronové sítě jsou schopny zpracovat relativně velká množství dat, mají schopnost generalizace - tedy zobecnění naučených vztahů - a v neposlední řadě, vzhledem k výpočetnímu aparátu, kterým jsou jednotlivé neurony tvořeny, mohou mít tyto sítě určitou odolnost vůči chybě, resp. šumu, který se na vstupu objevuje. Problémem ale je fakt, že v komplexních závislostech může být určitý šum obsažen i na výstupu těchto sítí. Vlastnosti neuronových sítí tedy svými vlastnosti řeší množství očekávaných problémů ve zkoumané oblasti.

2.3 Predikce

Predikcí je považován přístup, kdy na základě předchozího vývoje odhadujeme další vývoj veličiny. Tento vývoj ale může být závislý i na dalších parametrech,

kteře mohou do predikce více nebo méně zasahovat. Vliv těchto parametrů nemusí být jednoduše kvantifikovatelný.

2.3.1 Nelineární regrese a predikce časových řad

Nelineární regrese je podle [8] přístup, kdy se snažíme proložit nelineární funkcí množinu určitých dat tak, aby byla minimalizována interpolační chyba této funkce. Pro regresi lze použít celou řadu funkcí, ale pro experimentální data nelze předem určit typ a řád této funkce. Vzhledem k rozměru experimentálních dat není ani prakticky možné stavový prostor přehledně znázornit. Pro využití nelineární regrese by bylo zapotřebí postupně otestovat vytipované regresní funkce, které mají možnost aproximovat tak složitou závislost jako jsou meteorologické vlivy. Statistická predikce časových řad je komplexní statistický obor, typickou metodou pro použití statistiky na zadaný problém by mohla být dekompoziční metoda predikce časových řad [9]. Je zřejmé, že atmosferické vlivy lze dekomponovat na vlivy způsobené ročním obdobím, střídáním dne a noci a pak atmosferickými vlivy, které mají velkou dynamiku změn. Dlouhodobé odchylky lze aproximovat regresními metodami. Problém ale budou způsobovat náhlé změny v meteorologických datech. Ty jsou statisticky postihnuteľné například klouzavým průměrem a podobnými krátkodobými aproximacemi. I přes tyto možnosti ale nelze do statistik implementovat velmi rychlé dynamické změny a krátkodobé úniky způsobené turbulencemi, které se často vyskytují například u FSO spojů. Z popisu metod je zřejmé, že nevýhodou použití této metody je nutná znalost alespoň nějakých vlastností experimentálních dat. V případě jejich neznalosti pak statistická metoda nemůže poskytovat dostatečné univerzální výsledky.

2.3.2 Neparametrický data-fitting

Tento přístup lze podle [10] použít pro případy naměřených dat (tedy zatížených nejistotou), u kterých neznáme předem předpokládaný tvar funkce, kterou bychom těmito body mohli proložit (to je typický případ využití curve-fittingu). Účelem této metody je ale ve obvykle nalézt trend vývoje konkrétní množiny bodů. To není pro náš případ optimální, protože některé atmosferické vlivy mohou mít rychlou dynamiku změn a tímto přístupem bychom se o tyto změny mohli připravit. Tuto metodu je možné využít pro predikci dlouhodobého vývoje ale i jako doplněk nebo náhradu některé ze statistických metod z předchozího odstavce.

2.3.3 Neuronové sítě pro predikci

Neuronové sítě opět velmi dobře korespondují s řešeným problémem. Lze nalézt jejich využití v predikci časových řad (jsou známy jejich aplikace ve finančnictví, technologických procesech a v neposlední řadě i pro enviromentální vlivy). Kon-

krétní aplikace neuronových prediktorů v oblasti radiových a optických spojů ale zatím nebyla nikde konkrétně analyzována. Pomocí některých architektur neuronových sítí lze předpokládat dosažení dobrého časového prediktoru pro konkrétní environmentální vlivy na telekomunikační spoje.

2.4 Výběr vhodných metod

Z předchozí rešerše je na základě popsaných vlastností jednotlivých přístupů vybrána následující skupina metod, které nejlépe korespondují se zadanými problémy a metodika jejich použití bude podrobněji popsána dále v práci:

1. Analýza dat:

- Statistická analýza naměřených dat - zjištění statistického rozložení, korelace jednotlivých měřených veličin
- Analýza pomocí neuronových sítí - Shluková analýza pomocí Kohonenových samoorganizujících map, analýza závislostí jednotlivých měřených veličin

2. Klasifikační problémy:

- Neuronové sítě - typickou architekturou jsou vícevrstvé dopředné neuronové sítě, u kterých bylo dokázáno, že mohou fungovat jako univerzální a nelineární klasifikátor a aproximátor.

3. Prediktivní problémy:

- Neuronové sítě - dokáží díky svým schopnostem generalizace a odolnosti vůči chybám ve vstupních datech aproximovat obecné nelineární funkce. Jejich uplatnění se rovněž dá nalézt v predikci časových řad (typicky i finančních a technologických dat, která jsou ovlivňována vnějším prostředím a toto prostředí není možné do všech detailů popsat). V zadaném problému, kde jsou k dispozici relativně komplexní meteorologická data a kdy je cílem na kratší časový úsek odhadovat další vývoj kvality FSO spoje, se neuronové sítě používané pro predikce časových řad zdají být velmi vhodným kandidátem na aplikaci.

Kapitola 3

Současný stav problematiky

Tato kapitola popisuje aktuální stav problematiky metod odhadu a predikce atmosferických vlivů na radiové i optické bezdrátové spoje. V jednotlivých podkapitolách jsou rovněž popsány fyzikální principy, které uvedené telekomunikační spoje ovlivňují a je na ně odkazováno v následujících kapitolách. Popsané hlavní vlivy a jejich vlastnosti jsou uvažovány při tvorbě metodiky, která je cílem práce.

3.1 Oblast radiových spojů

Pro řešení vlivu atmosferických vlivů na radiové spoje v současnosti existuje řada doporučení ITU-R [11, 12, 13, 14, 15, 16]. Jejich obsahem jsou především postupy pro výpočet útlumu způsobeného hydrometeory, vzdušnými aerosoly a podobně.

Pro výpočet šíření v závislosti na stavu atmosféry existuje řada modelů, typicky parabolická rovnice. K jejímu řešení existuje celá řada algoritmů a přístupů publikovaných například v [17], [18], [19] a řadě dalších.

Další podrobnosti k jednotlivým vlivům prostředí a výpočetním modelům jsou podrobněji popsány dále v textu.

První část této kapitoly se okrajově dotýká dálkového průzkumu Země. Ohledně radiového průzkumu troposférických jevů, které mají vliv na šíření elektromagnetické vlny byly publikovány například práce [20] a [21]. Tyto práce se věnují především refraktivitě v souvislosti se vzdušnými aerosoly.

Byly publikovány i články zaměřené na radiovou analýzu vlastností hydrometeorů [22]. Existují i další metody pro odhad profilu refraktivity na experimentálním spoji využitého v této práci [23].

3.1.1 Publikované metody pro odhad refraktivity v troposféře

Znalost změn refraktivity v čase a prostoru je důležitá pro pochopení mechanismu šíření elektromagnetické vlny v prostředí a pro zjištění jakým způsobem vznikají úniky. Refraktivita silně ovlivňuje šíření v okamžiku, kdy šíření není silně ovliv-

ňováno silnějšími vlivy jako je absorpce a rozptyl na hydrometeorech. Byly prováděny experimenty zaměřené především na vznik vlnovodných kanálů a to především v přímořských oblastech, kde jsou povětrnostně naprosto jiné podmínky než ve vnitrozemí. Tyto experimenty byly publikovány v literatuře - např. v [24] a [25]. Podmínky, jaké panují mezi vysílačem a přijímačem, byly v těchto experimentech odvozovány právě na základě naměřených úrovní mikrovlnných spojů. Naproti tomu například v publikaci [26] je výskyt vlnovodného kanálu predikován pouze na základě meteorologických měření. Existuje pouze několik publikací, které by se zabývaly kombinovanými experimenty jako je ten použitý v této práci. Kombinované experimenty lze najít v [27], [28], [29], [30]. Rovněž existují i publikace, které pracují s refraktivitou v troposféře za pomoci různých neuronových sítí. Tyto experimenty opět nezahrnují kombinaci meteorologických a radiových dat - viz. například [31]. Dalším přístupem k odhadu profilu refraktivity za pomoci neuronových sítí je využití jiných spojů než terestriálních. Publikace [32] popisuje odhad refraktivity atmosféry pomocí naměřených dat z radiových spojů k GPS družicím. Na experimentálním spoji, který byl využit i v této práci, byla v [33] publikována metoda, která na základě známých fyzikálních vztahů predikuje gradient refraktivity za využití překážky v trase spoje. Gradient refraktivity je touto metodou odhadován až s velmi dobrou 70% přesností.

3.2 Šíření radiového signálu v troposféře

Ačkoli předmětem této práce není výzkum a kvantifikace konkrétních fyzikálních zákonitostí vlivu atmosféry na bezdrátové spoje, je třeba popsat meteorologické jevy a jejich vlastnosti, které tyto spoje ovlivňují. Z následujícího přehledu jednotlivých atmosferických jevů lze předpokládat multidimenzionalitu a dynamiku jejich vlivu na komunikační spoje.

3.2.1 Meteorologické jevy ovlivňující šíření

3.2.1.1 Srážky

Srážky mohou mít různou intenzitu a mohou být tvořeny jak vodními kapkami, tak i sněhovými vločkami nebo ledovými krystalky. Měření srážek lze provádět pomocí sítě srážkoměrů nebo pomocí meteorologického radaru. Srážkoměry poskytují větší přesnost měření srážek, ale radar je schopen lépe pokrýt sledované území, kdy je měřena odrazivost a množství (distribuce) srážek je z ní přepočítáváno pomocí empirických vztahů, kterých existuje celá řada pro různé druhy atmosfér a z velké části jsou založeny na práci [34] a [35]. Z hlediska šíření elektromagnetické vlny se kapka chová jako ztrátové dielektrikum, na kterém dochází k absorpci energie. V kapce se indukují posuvné proudy a elektromagnetická energie se přeměňuje na tepelnou. Zároveň na kapce dochází k rozptylu, který je podrobněji popsán v kapitole 3.5.2.

3.2.1.2 Přizemní oblačnost

Typickým jevem, který může výrazně ovlivňovat šíření je přizemní oblačnost a její typický druh Stratus ([36]). Tato oblačnost se vyskytuje v nejnižších vrstvách troposféry, ve kterých již v některých případech prochází elektromagnetická vlna z pozemních spojů. Vzniká velmi často z mlhy, která se zvedá od země a vytváří tak jednolitě vrstvy, které mají malou vertikální mohutnost, ale jsou rozprostřeny ve velké ploše. Často má vzhled jednolitě šedé plochy přes celou oblohu. Dále může Stratus vznikat při odparu z ploch zadržujících vzdušnou vlhkost. V tom okamžiku mívá Stratus vzhled menších roztrhaných útvarů, které začínají na úrovni korun stromů a mohou být ve více vrstvách. Stratus je vodním oblakem a v zimních měsících může obsahovat i krystaly ledu. Nezpůsobuje výrazné srážky, ale spíše jemné mrholení a v zimě vypadávání sněhových krupiček. Stratus je svojí strukturou velmi podobný mlze, ale meteorologicky se jedná o jiný jev. Výpočet útlumu při průchodu nízkou oblačností je možno nalézt v [14].

3.2.1.3 Mlha

Mlha obvykle leží ve výškách od zemského povrchu nejvýše do stovek metrů. Vzniká při teplotách kolem rosného bodu, kdy se vzduch nasytí vodními parami. Je tedy tvořena vodními kapkami, případně v zimě může být tvořena ledovými krystalky. Mlhy mohou být způsobeny různými faktory. V blízkosti měst a průmyslových podniků lze identifikovat mlhy s obsahem zplodin jako jsou například produkty hoření, dále mohou být mlhy způsobeny při pohybu vlhkého vzduchu v údolních kotlinách nebo třeba při pohybu teplé masy vzduchu na chladnější zemský povrch. Efektem mlhy je tedy výrazné snížení dohlednosti od řádu desítek metrů do cca 5 km. Vzhledem k principu vzniku je výskyt mlh časově závislý a často se vyskytují při inverzním charakteru počasí. Mlha, podobně jako v případě srážek, způsobuje ztráty kvůli rozptylu a absorpci na částicích obsažených v atmosféře. Vliv mlhy je markantní až při vyšších radiových frekvencích a v optickém spektru.

3.2.1.4 Ostatní jevy

Vliv na šíření radiového signálu mají další meteorologické jevy - od větru, se kterým se přesouvá oblačnost a mlhy, ochlazuje zem a zvedá prach z polí, až po změny teplot a tlaků na frontálních rozhraních. Vliv těchto jevů je prakticky těžko kvantifikovatelný, protože tyto jevy obvykle působí v kombinaci a nelze tak od sebe oddělit vliv větru a oblačnosti, kombinace mlhy a větru a podobně. Dalším problémem je určování těchto vlivů na delší spoje, protože obvykle nelze změřit meteorologické veličiny s dostatečnou jemností po celé délce spoje.

3.2.2 Útlum způsobený průchodem elektromagnetické vlny prostředím

Doporučení ITU-R P.676 [12] zahrnuje metodiku, pomocí které je možné určit specifické útlumy atmosferických plynů pro frekvence do 1000 GHz. Tento útlum je dán především rezonancemi molekul plynů obsažených v atmosféře. Vstupními daty pro tyto metodiky je především teplota, tlak a vlhkost. Jsou zde přehledně uvedeny průběhy útlumu pro různé frekvence pro atmosféru s definovanou koncentrací vodních par. Zároveň je zde uveden postup výpočtu specifického útlumu pro různé nadmořské výšky.

3.2.3 Útlum hydrometeory

Hydrometeory jsou všechny meteorologické jevy, které v atmosféře udržují větší množství vody v libovolném skupenství s výjimkou oblačnosti. Typicky se jedná o srážky všech druhů, mlhy a zvířený sníh. Útlum hydrometeory se výrazně projevuje od frekvence 10 GHz a výš. Obecně se k řešení útlumu elektromagnetického vlnění způsobeného hydrometeory přistupuje statisticky. Tato statistika může být založená buď na plošných vlastnostech deště nebo na analýze vlastností vodních kapek (nebo jiných vodních objektů), které se v daném hydrometeoru vyskytují. Specifický útlum způsobený deštěm je úměrný intenzitě srážek. Základní používaný model uvádí doporučení ITU [11], který uvádí možnosti výpočtu útlumu spoje na definované frekvenci, polarizaci, délce spoje a především statisticky definovaných intenzitách deště v průměrném roce. Tento model je platný pro spoje o frekvencích do 50 GHz a délce do 60 km. Zpráva [37] uvádí přehled a srovnání jednotlivých modelů odhadu intenzity srážek pro průměrný rok, průměrný nejhorší měsíc a pro kombinaci průměrného roku a průměrného nejhoršího roku. Metoda ITU-R [11] je hodnocena společně s metodami Lefrancois [38], Stutzman-Dishman [39] a Lin I [40] jako nejspolehlivější pro testovací spoje rozmístěné po celé Evropě.

3.2.4 Atmosférické turbulence

V atmosféře probíhají relativně chaotické toky plynů, toto proudění může být laminární, ale při dosažení vyššího Reynoldsova čísla se mění na turbulentní. Při turbulentním proudění vzduchu vznikají víry, které způsobují homogenizaci vlastností atmosféry v dané malé oblasti. Víry mohou mít různé velikosti a jejich tvar a velikost není stálá. Víry se s postupem času rozpadají na menší. Turbulence mohou mít velikosti od řádu milimetrů až do desítek metrů. V těchto oblastech se oproti okolí mění rychlost a směr proudění větru, teplota, vlhkost a především i refraktivita prostředí. Turbulence se často vyskytují i v místech konvektivního proudění vzduchu. Průběh a vlivy vzdušných turbulencí je podrobněji popsán v knize [41].

Hlavní vliv turbulencí na šíření elektromagnetické vlny jsou změny v refraktivitě, jejíž dopady budou rozebrány v dalším odstavci. Důsledkem průchodu elektromagnetické vlny je rovněž scintilace, což je časově proměnná změna v úrovni přijímaného signálu. Tato scintilace bývá obvykle počítána jako rozptyl úrovně signálu v určitém časovém úseku. Dalším efektem turbulencí je rozptyl, který nastává na malých nehomogenitách, kterými mohou být i turbulence. Míra rozptylu je závislá na frekvenci vlnění a velikosti nehomogenit.

3.2.5 Troposférická refrakce

K troposférické refrakci dochází při průchodu elektromagnetické vlny přes prostředí s různým indexem lomu. Literatura [42] uvádí index lomu troposféry v rozmezí cca od 1,000110 do 1,000325. Prakticky se ale pro vyjádření indexu lomu zavádí refraktivita N , která má jednotku N-unit a je definována:

$$N = (n - 1) \cdot 10^6, \quad (3.1)$$

kde n je index lomu.

Výpočet refraktivity uvádí doporučení ITU [43], které zároveň uvádí typické hodnoty refraktivity pro celou Evropu a svět a vztahy pro přepočtení refraktivity v závislosti na nadmořské výšce. Refraktivita je definována vztahem

$$N = \frac{77,6}{T} \left(P + 4810 \frac{e}{T} \right), \quad (3.2)$$

kde T je absolutní teplota [K], P je atmosferický tlak [hPa] a e je tlak vodních par:

$$e = \frac{H \cdot e_s}{100}, \quad (3.3)$$

kde H je relativní vlhkost [%] a e_s je tlak nasycených vodních par při příslušné teplotě. Tlak vodních par lze ještě získat za použití ITU-R P.836 [13], kde lze v příloze najít parametr hustoty vodních par ρ , který je použit ve vztahu:

$$e = \frac{\rho \cdot T}{216,7}. \quad (3.4)$$

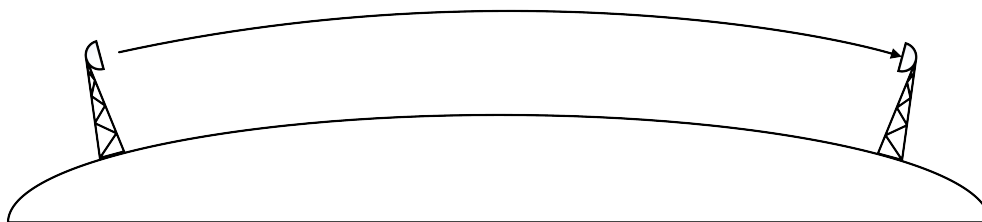
V praktických aplikacích se obvykle sleduje výšková změna refraktivity, resp. gradient refraktivity v určité vrstvě troposféry. V nejnižších vrstvách troposféry, kterými se zabývá tato práce je podle [43] výškový gradient refraktivity:

$$\frac{dN}{dh} = -40N/km. \quad (3.5)$$

Tato hodnota je statisticky zjištěná hodnota. Prakticky se ale gradient refraktivity velmi silně mění se změnami počasí, ale i se střídáním dne a noci a ročních období. Hodnota gradientu ale vyjadřuje průměrnou hodnotu v několika kilometrové vrstvě nejbliž k zemi. Pokud budeme průběh refraktivity zjišťovat v menších

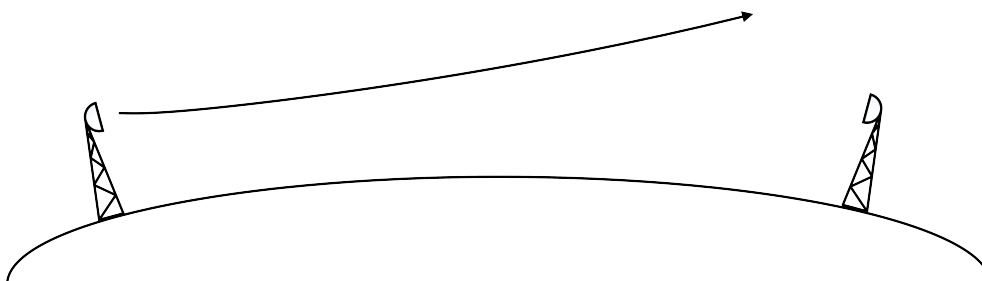
krocích, zjistíme, že průběh refraktivity rozhodně není lineární, ale lze ho po určitých krocích lineárně interpolovat. Tím lze získat výškový profil refraktivity.

Pokud elektromagnetická vlna prochází různými vrstvami s různou refraktivitou, dochází k jeho zakřivení - to je dáno Snellovým zákonem lomu. Literatura [42] uvádí výpočet poloměru zakřivení takového paprsku v závislosti na gradientu refraktivity. Pro hodnotu gradientu refraktivity -40 N/km odpovídá poloměru zakřivení cca 25000 km . To znamená, že se vlna nešíří přímočaře, ale ohýbá se směrem k zemi. Tomuto zakřivení tedy odpovídá standardní refrakce - zobrazena na obr. 3.1.



Obr. 3.1: Standardní refrakce

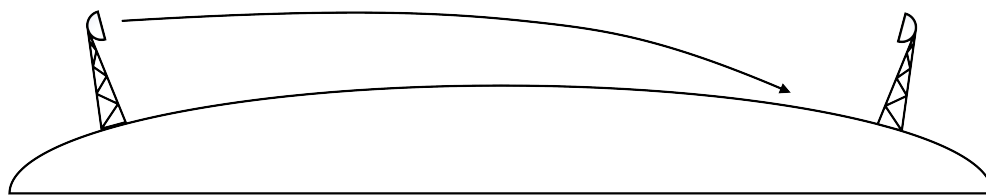
Dalším efektem, který může nastat, je subrefrakce (obr. 3.2). Ta nastává v okamžiku, kdy je paprsek procházející troposférou zakřiven s větším poloměrem než při standardní refrakci. Této situaci odpovídá gradient refraktivity $dN/dh > -40 \text{ N/km}$. Může nastat i situace, kdy se paprsek odklání od země, tomu odpovídá $dN/dh > 0 \text{ N/km}$.



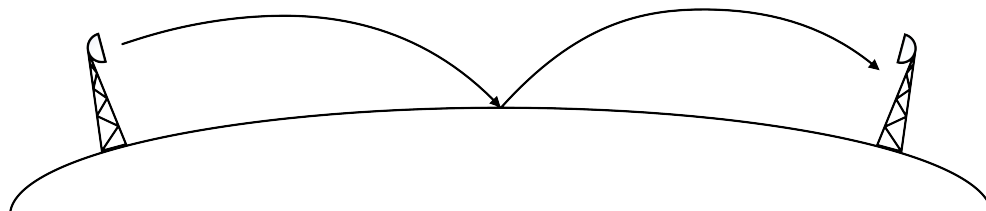
Obr. 3.2: Subrefrakce

Superrefrakce je dalším jevem, který nastává v atmosféře. Je zobrazen na obr. 3.3 a jedná se o ohyb paprsku, který je s menším poloměrem než v případě standardní refrakce, ale zároveň větším než je poloměr Země. Pro výškový gradient refraktivity tedy platí $dN/dh < -40 \text{ N/km}$.

V okamžiku, kdy je poloměr ohybu paprsku shodný s poloměrem Země, nastává případ, kdy by se paprsek mohl šířit rovnoběžně se zemí. Tomu odpovídá gradient refraktivity $dN/dh = -157 \text{ N/km}$. Pro případ, kdy se paprsek ohýbá s poloměrem menším než poloměr Země, dopadá elektromagnetická vlna na zem a odráží se od ní. Míra tohoto odrazu je dána i parametry země v místě dopadu (typicky její vodivostí). Tento jev se nazývá přízemní vlnovodný kanál (surface duct) a je naznačen na obr. 3.4

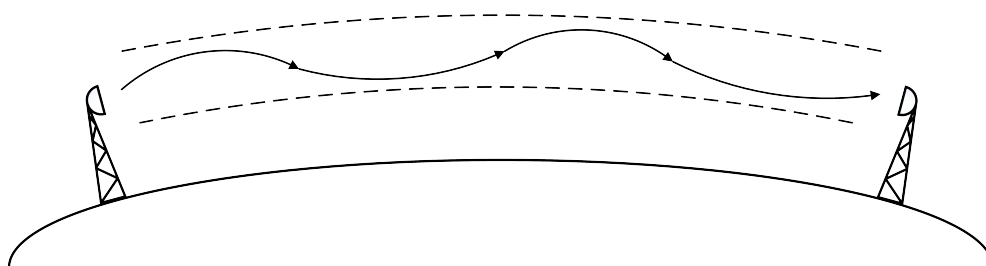


Obr. 3.3: Superrefrakce



Obr. 3.4: Přízemní vlnovodný kanál

Speciální případ vlnovodného kanálu nastává, pokud jsou v určité výšce nad terénem splněny předpoklady pro vznik vlnovodného kanálu. Pak při vhodné elevaci paprsku může docházet ke vzniku vyvýšeného vlnovodného kanálu, jehož příklad je uveden na obr. 3.5.



Obr. 3.5: Vyvýšený vlnovodný kanál

3.3 Metody výpočtu vlivu troposférické refrakce na radiové spoje

Ze statistických hodnot refraktivity uvedených v [43] lze získat pouze statistické hodnoty úniků v určitých časových intervalech. Pro případy řešení konkrétního vlivu rozložení refraktivity existuje možnost numerického řešení parabolické rovnice (3.6) publikované v [19] nebo některé z metod geometrické optiky jako je raytracing. Parabolická rovnice je definována jako:

$$\frac{\partial^2 E}{\partial h^2} - 2j \left(\frac{2\pi}{\lambda} \right) \frac{\partial E}{\partial d} + \left(\frac{2\pi}{\lambda} \right)^2 (n^2 - 1)E = 0, \quad (3.6)$$

kde E je intenzita elektrického pole, n je index lomu, d vzdálenost, h je výška a λ vlnová délka. Její řešení je relativně náročné a bylo už publikováno množství

článků zabývajících se konkrétními řešeními některých problémů šíření.

Raytracing v oblasti šíření je metoda, kdy se v určitém kroku studuje průběh jednotlivých paprsků a řeší se jeho zakřivení na základě průběhu refraktivity a tedy i efektů uvedených v odstavci 3.2.5. Jeho srovnání s řešením parabolické rovnice je uvedeno například v [44] a [45].

3.4 Oblast bezdrátových optických spojů

V oblasti atmosferických vlivů na FSO spoje byla publikována řada prací, která se zabývá odhady specifických útlumů atmosferických jevů jako je mlha, déšť a sněh. Tyto specifické útlumy jsou převážně založeny na znalosti statistických vlastností příslušných atmosferických jevů.

Pro útlumy způsobené mlhou byly publikovány například Kimův model [46] a Kruseův model [47], které jsou závislé na viditelnosti, vlnové délce a množství vodních kapek v prostředí. Byly publikovány i srovnání jednotlivých metod v [48].

Pro specifický útlum způsobený deštěm existují doporučení ITU-R [16].

Dalšími studovanými vlivy jsou scintilace optických spojů. Pro moderní komunikační služby je vyžadována velká šířka pásma, a zároveň musí být definovaná minimální šířka pásma, aby byla zaručena bezchybnost přenosu dat v reálném čase. Právě scintilace způsobují zvýšení chybovosti linky a tím zmenšení šířky pásma. Scintilace jsou rovněž spojeny s turbulencemi v přenosovém médiu. Studium scintilací způsobených turbulencemi se zabývají například publikace [49], [50], [51], [52] a [41].

Vzhledem k tomu, že FSO je už relativně dlouho používaná technologie, jsou publikovány výsledky dlouhodobých měření spolehlivosti těchto spojů například [53] a [54].

V souvislosti s FSO spoji se objevují studie diverzitních spojů. V oblasti zájmu je studium prostorová diverzity optických spojů a pak statistické zpracování spolehlivosti hybridních spojů, ve kterých se k přenosu dat používá kombinace FSO a radiového spoje. Zde je předpoklad, že průběh atmosferických vlivů na FSO a radiové spoje není shodný, tudíž lze u těchto spojů odhadovat diverzitní zisk, závislý na prostředí. Různé technologie těchto hybridních spojů byly publikovány například v [55], [56] a [57]. Studie zabývající se spolehlivostí takových spojů byly publikovány v [58], [59] a dalších.

3.5 Vlivy prostředí na FSO spoje

Vliv počasí na funkci FSO spojů je důležitým faktorem, který musí být při návrhu ([60]) těchto spojů brán v úvahu. Při přenosu dat pomocí těchto spojů se jedná vlastně o přenos elektromagnetické energie a tento přenos je narušován vlivy jako je rozptyl na částicích, které se mohou ve vzduchu objevovat. Na těmito spoji používaných vlnových délkách se jedná zejména o vzdušné aerosoly,

s velikostí části řádově od $0,1\mu m$ do $100\mu m$ (to může být například i mlha, ale i další látky popsané dále v této kapitole), které způsobují velké ztráty při šíření elektromagnetické vlny. Tyto ztráty jsou do velké míry způsobeny Mieovým rozptylem. Toto tvrzení je podpořeno teoretickými studiemi [61], [62] a byly publikovány i výsledky měření jako například [63] a [64].

Dalším nezanedbatelným vlivem jsou tepelné turbulence. Tyto turbulence mohou způsobovat vychýlení a rozostření světelného svazku. Teoreticky byly vlivy turbulencí rozebrány v [65], [49] a experimentálně ověřené vlivy byly publikovány například v [66].

3.5.1 Extinkce

Extinkce je způsobena dvěma hlavními vlivy: absorpcí a rozptylem v médiu, tedy atmosféře. Tyto jevy odebírají laserovému paprsku energii a tím omezují dosah FSO spojů. Extinkci v principu způsobují samotné atomy a molekuly média a zároveň i větší částice - vzdušné aerosoly. Jak je popsáno v [67], míra extinkce je přímo úměrná intenzitě záření a množství materiálu, kterým paprsek prochází za předpokladu konstantního složení média a fyzikálních vlastností jako je vlhkost, teplota a podobně. Vliv extinkce je na intenzitu procházejícího záření popsán rovnicí

$$I_\nu = I_0 e^{-\tau_\nu}, \quad (3.7)$$

kde τ_ν je optická tloušťka prostředí pro záření o vlnové délce μ . Optická tloušťka prostředí mezi body x_1 a x_2 je definována jako

$$\tau_\nu(x_1, x_2) = \int_{x_1}^{x_2} \beta_{ext,\nu}(s) ds, \quad (3.8)$$

kde $\beta_{ext,\nu}(s)$ je koeficient extinkce závislý na vlnové délce a na samotném prostředí. Tento koeficient je součtem koeficientu absorpce a rozptylu prostředí.

Hlavním přispěvatelem k útlumu paprsku o vlnových délkách používaných v FSO spojích je absorpce na molekulách média. Tato absorpce způsobuje změny elektronových stavů na molekulách chemických sloučenin, kterými paprsek prochází. V souvislosti s tím je zřejmé, že absorpce bude vždy závislá na použité vlnové délce. Pro vlivy molekulární absorpce a souvisejících jevů existují modely, které dokáží odhadnout koeficient extinkce v závislosti na viditelnosti a vlastnostech hydrometeoru, které jsou ve směru šíření paprsku v médiu. Těchto statistických modelů existuje celá řada, např. [63]. Většina těchto modelů je založena na statistických vlastnostech hydrometeorů. Tato statistika je však obtížně měřitelná běžnými meteorologickými prostředky.

Dalším vlivem může být extinkce na větších částicích, typicky vzdušných aerosolech. Tyto částice mohou být zastoupeny v atmosféře po relativně dlouhou periodu. Jejich velikost uvažujeme řádově od jednotek nm do desítek μm a typickými zástupci jsou z přírodních zdrojů:

- prach z půdy a skal, vulkanický prach,
- popílky,
- částice navázané na přírodní výrony plynů jako je H_2S , NH_3 , NO_8 a uhlovodíků, které uvolňují rostliny,
- v přímořských oblastech i uvolněná mořská sůl

Z člověkem vytvořených částic jsou to především:

- přímé emise - průmyslové, lokální topení apod.,
- částice navázané na vypouštěné emise SO_2 , NO_8 a další

Životnost výše uvedených částic v atmosféře je do velké míry závislá na jejich velikosti. Na základě práce [68] lze říci, že nejmenší částice (pod 100 nm) mohou v atmosféře přetrvávat roky a větší částice (nad 10 μm) řádově hodiny až minuty. Samozřejmostí je i rozdílná koncentrace těchto částic v různých prostředích (v městských prostředích jsou více přítomny průmyslové emise a ve venkovském prostředí bude převládat prach ze zemědělské půdy a obecně koncentrace cizích látek bude nižší než ve městě). V [69] byl publikován přehled výpočetních a měřících metod pro analýzu absorpce světla na vzdušných aerosolech.

3.5.2 Rozptyl

Rozptyl světla nastává v okamžiku, kdy elektromagnetická vlna dopadá na překážku nebo nehomogenitu. Tato vlna interaguje s diskretní částicí a v jednotlivých molekulách vzniká indukovaný dipólový moment, který je sám zdrojem elektromagnetického vlnění. Většina záření je emitována se stejnou frekvencí jako dopadající světlo, ale toto záření je emitováno v různých směrech. Lze tedy říci, že rozptyl světla není dán jen odrazem fotonů nebo elektromagnetické vlny od povrchu objektu, ale komplexní interakcí molekulární struktury překážky a dopadající vlny. Pro kvantifikaci jevu rozptylu světla jsou v současnosti nejpoužívanější dva přístupy - Rayleighův rozptyl a Mieův rozptyl.

3.5.2.1 Rayleighův rozptyl

Rayleighův rozptyl, je původně formulován pro malé dielektrické (neabsorbující) kulové částice. Kritériem pro využití Rayleighova rozptylu je $\alpha \ll 1$, kde

$$\alpha = \frac{2\pi r}{\lambda}. \quad (3.9)$$

Parametr λ je vlnová délka dopadajícího polarizovaného světla a r je poloměr částice, jejíž rozptyl analyzujeme. V práci [70] lze najít odvození intenzity rozptýleného světla pro absorbující i neabsorbující částice. Lze říci, že pro $m = n - ik$,

kde n je index lomu materiálu částice způsobující rozptyl a k je materiálová konstanta, která má vztah k absorpci částice. Účinné průřezy (efektivní odrazné plochy) pro takovou částici lze vyjádřit jako

$$\sigma = \sigma_{\text{rozptyl}} + \sigma_{\text{absorpce}}, \quad (3.10)$$

kde jednotlivé součásti lze vyjádřit:

$$\sigma_{\text{rozptyl}} = \frac{2\lambda^2\alpha^6}{3\pi} \left| \frac{\bar{m}^2 - 1}{\bar{m}^2 + 2} \right|^2, \quad (3.11)$$

$$\sigma_{\text{absorpce}} = \frac{-\lambda^2\alpha^3}{\pi} \text{Im} \left(\frac{\bar{m}^2 - 1}{\bar{m}^2 + 2} \right). \quad (3.12)$$

Složka σ_{absorpce} může být zanedbána pro dielektrické částice (tj. když $k = 0$).

Ztráta energie dopadajícího záření s intenzitou I_0 způsobená rozptylem bude vyjádřena jako:

$$E = I_0\sigma. \quad (3.13)$$

Rayleighův rozptyl je přes svoje omezení relativně často využíván. V případech, kdy je třeba použít komplexnější modely rozptylu, lze využívat Mieův rozptyl. Nevýhoda tohoto přístupu je nemožnost použití tohoto přístupu pro částice, které mají srovnatelnou velikost s vlnovou délkou.

3.5.2.2 Mieův rozptyl

Jak uvádí literatura [71], Mieův rozptyl je metoda, jejíž myšlenky byly poprvé publikovány Gustavem Miem již v roce 1918, kdy byl nastíněn postup výpočtu rozptylu světla pomocí Maxwellových rovnic. Mieův rozptyl je tedy ucelená teorie, která je schopna vysvětlit a v mnoha případech i vypočítat rozptyl světla na částicích, pro které neplatí Rayleighův přístup. Pomocí tohoto přístupu lze vypočítat rozptyl na částicích, jejichž rozměr je srovnatelný nebo větší, než je vlnová délka dopadajícího záření. Analogicky při zachování velikosti částice lze studovat rozptyl pro vyšší frekvence. Pro Mieovu teorii byly publikovány metody, které umožňují výpočet rozptylu na nesférických částicích - přehled uveden například v [72]. Obecnou nevýhodou Mieova přístupu k rozptylu je jeho výpočetní náročnost. Z toho důvodu bývá Rayleighova teorie upřednostňována v případech, ve kterých je použitelná, protože rovněž poskytuje správné výsledky. Pro představu složitosti výpočtu zde uvádím vztahy pro výpočet účinných průřezů σ_{rozptyl} a σ_{absorpce} , stejně jako u Rayleighova rozptylu v předchozím odstavci:

$$\sigma_{\text{rozptyl}} = \frac{\lambda^2}{2\pi} \sum_{n=0}^{\infty} \left((2n+1)(|a_n|^2 + |b_n|^2) \right) \quad (3.14)$$

$$\sigma_{\text{absorpce}} = \frac{\lambda^2}{2\pi} \sum_{n=0}^{\infty} \left((2n+1)\text{Re}(a_n + b_n) \right). \quad (3.15)$$

Parametry a_n a b_n jsou definovány za použití sférické Besselovy funkce prvního řádu $J_{n+1/2}(z)$ a sférické Hankelovy funkce druhého řádu $H_{n+1/2}(z)$:

$$a_n = \frac{\Psi_n(\alpha)\Psi'_n(m\alpha) - m\Psi_n(m\alpha)\Psi'_n(\alpha)}{\xi_n(\alpha)\Psi'_n(m\alpha) - m\Psi_n(m\alpha)\xi'_n(\alpha)} \quad (3.16)$$

$$b_n = \frac{m\Psi_n(\alpha)\Psi'_n(m\alpha) - \Psi_n(m\alpha)\Psi'_n(\alpha)}{m\xi_n(\alpha)\Psi'_n(m\alpha) - \Psi_n(m\alpha)\xi'_n(\alpha)} \quad (3.17)$$

$$\Psi_n(z) = \left(\frac{\pi z}{2}\right)^{\frac{1}{2}} J_{n+1/2}(z) \quad (3.18)$$

$$\xi_n(z) = \left(\frac{\pi z}{2}\right)^{\frac{1}{2}} H_{n+1/2}(z) \quad (3.19)$$

Na numerické řešení problémů pomocí Mieovy teorie rozptylu světla existuje celá řada numerických metod, jejich přehled a srovnání lze najít například v [71].

3.5.3 Turbulence

Turbulence, jak byly popsány v kapitole 3.2.4, se uplatňují i u optických spojů FSO.

V literatuře [73], [52] a [74] lze najít naměřené charakteristiky scintilací FSO spojů, kdy míra scintilace je úměrná síle turbulencí a délka úniků a výrazných scintilací má trvání od jednotek milisekund. Délka a míra scintilací se mění v závislosti na délce spoje, počasí, ročním období a dokonce i v průběhu dne a míře slunečního svitu. Střední hodnota scintilací dosahuje až jednotek dB na kilometrové vzdálenosti.

3.6 Zhodnocení současného stavu

Současný stav znalostí vlivu meteorologických jevů je již velmi dobře popsán celou řadou modelů. Tyto modely jsou ale definované především statisticky a nejsou primárně závislé na čase nebo na kombinaci jednotlivých atmosferických fenoménů. Tato práce se tedy zaměří na kvantifikaci kombinace více současně působících jevů (tak jak je to v atmosféře běžné) a působení kombinace těchto vlivů v závislosti na čase. Zároveň bude snaha co nejméně využívat informace z existujících modelů.

Kapitola 4

Cíle práce

Základním cílem práce je ověřit, je-li možné predikovat chování bezdrátových spojů v atmosféře bez znalosti fyzikálně vyjádřených vlivů jednotlivých meteorologických veličin na radiový nebo optický signál. Na základě takto obecně definovaného cíle je stanoveno několik dílčích cílů, které budou postupně řešeny a vyhodnocovány. Jednotlivé kroky vedou k sestavení jednotné metodiky zpracování dat pomocí metod umělé inteligence.

1. Výběr vhodného prostředku pro zpracování experimentálních dat.

Na základě výsledků rešerše metod uvedené v předchozí kapitole byla pro práci s experimentálními daty zvolena metoda neuronových sítí. Tato oblast je velmi rozsáhlá a v dalších kapitolách budou podrobně diskutovány jednotlivé architektury neuronových sítí použitelné pro řešení zadaných problémů. Tyto architektury budou postupně ověřeny na experimentálních datech a výkonnost bude analyzována.

2. Ověření vhodnosti zvolené výpočetní metody na problému detekce množství srážek v trase experimentálního spoje.

Pomocí jednodušších architektur neuronových sítí bude otestována jejich vhodnost pro danou problematiku. Vzhledem k novosti neuronového přístupu k datům z oblasti atmosferických vlivů na komunikační spoje je touto analýzou stanovena orientační přesnost odhadu, kterou dokáží neuronové sítě poskytnout.

3. Návrh metodiky pro odhad gradientu refraktivity pomocí neuronové sítě na základě dat naměřených na experimentálním spoji.

Cílem bude najít nový postup pro aplikaci neuronové sítě na odhad gradientu refraktivity v oblasti experimentálního spoje, který je již v literatuře relativně dobře popsán prostředky statistiky a existujícími modely. Tato metodika bude naprosto oproštěna od známých fyzikálních principů a využívat pouze naměřená data bez zakomponování jejich skutečného významu.

4. Návrh metodiky pro konstrukci neuronové sítě, která bude analyzovat vliv aktuálního počasí na kvalitu FSO spoje.

Metodika navržená pro radiové spoje poskytuje v některých případech relativně dobré přesnosti predikce, ale další zpřesnění není kvůli povaze experimentu už reálné. Protože není vzhledem k povaze experimentu možné překonat popsané problémy, je původní metodika použita na obdobný problém – odhad vlivu aktuální meteorologické situace na FSO spoje. Lze předpokládat, že vzhledem k dostupnosti kvalitních dat z experimentálních spojů, bude tato metodika poskytovat přesnější výstupy. Zároveň ji bude možno dále zpřesňovat – například dalšími meteorologickými čidly.

Kapitola 5

Neuronové sítě a jejich použití

Tato kapitola popisuje základní principy a terminologii neuronových sítí. Jejím účelem je uvést problematiku neuronových sítí a použitou terminologii, která není ve všech pramenech jednotná. Jednotlivé podrobnosti budou dále rozvíjeny ve vytvářené metodice v navazujících kapitolách této práce. Jsou zde rovněž uvedeny informace o současném stavu problematiky v oblasti neuronových sítí, jejichž typy jsou v této práci využívány.

5.1 Historie

Myšlenka neuronových sítí je z pohledu historie výpočetní techniky stará téměř stejně jako první generace počítačů. Vždy se lidé zamýšleli nad tím, jak funguje lidský mozek a vůbec celá lidská nervová soustava a jak lze znalosti této soustavy použít k vytvoření umělé inteligenci blízke té lidské. To se zcela jistě zatím nikomu nepovedlo, ale již desítky let vznikají pokusy napodobit lidský mozek počítačem nebo elektronikou. Počátky umělých neuronových sítí lze datovat do 40. let 20. století, kdy vznikl první McCulloch-Pittsův neuron (MCP neuron), který byl inspirován strukturou biologického neuronu, který je uveden na Obr. 5.1.

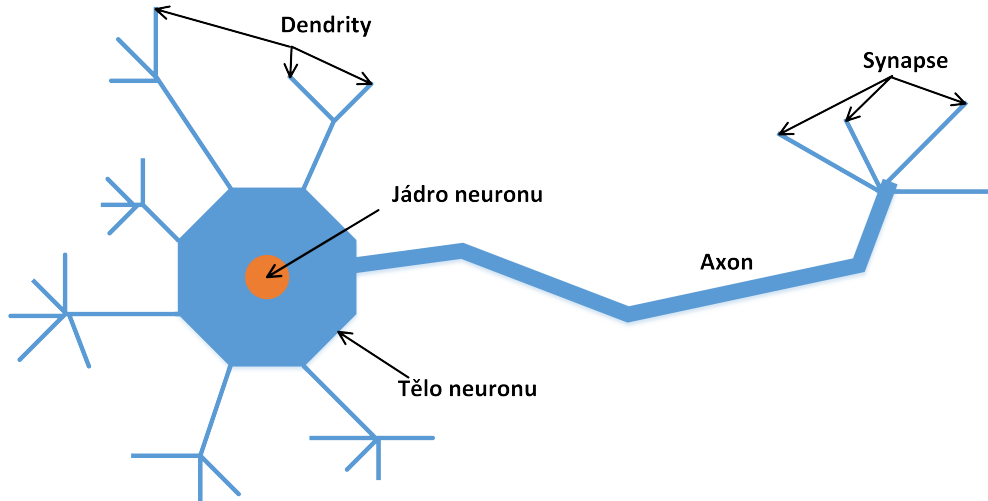
MCP neuron byla jen prahová logika, která poskytovala výstup 0 nebo 1 pro neuron s n vstupy:

$$y = \begin{cases} 0 & \text{pro } S < \theta \\ 1 & \text{pro } S \geq \theta \text{ a nesmí být aktivní žádná inhibující cesta} \end{cases}, \quad (5.1)$$

kde θ je prahová hodnota každého jednotlivého neuronu a S je vnitřní potenciál neuronu definovaný jako

$$S = \sum_{i=1}^n w_i x_i, \quad (5.2)$$

kde w_i je váha na každém vstupu neuronu, která může být kladná i záporná (pak je vstup nazýván jako excitující resp. inhibující) a x_i je vstupní hodnota.



Obr. 5.1: Biologický neuron, který se stal předlohou pro umělý neuron

MCP neuron mohl tedy dostávat vstupy od libovolného počtu jiných neuronů. Správným nastavením vah a prahů byl tento neuron schopen simulovat logické funkce jako AND, OR, NOT. Samotný MCP neuron nebyl schopen reprezentovat funkci XOR, ale v kombinaci s více neurony podle rozkladu:

$$x_1 \text{ XOR } x_2 = (x_1 \text{ OR } x_2) \text{ AND } \text{NOT} (x_1 \text{ AND } x_2) \quad (5.3)$$

Pomocí kombinace MCP neuronů je tedy možno vytvořit jakoukoli logickou funkci. Jejich problémem bylo to, že váhy a prahy byly pevné a nebylo možné zavést jakékoli učení.

V roce 1958 bylo publikováno vylepšení MCP neuronu, které se nazývá perceptron [75], případně podle autora Rosenblattův perceptron. Perceptron se proti MCP neuronu liší především v neexistenci absolutně inhibujících cest a v tom, že práh už není pevný, ale lze ho měnit pro každý neuron zvlášť stejně tak jako váhy jednotlivých vstupů perceptronu. Lze tedy říci, že rovnice (5.1) pro výstup neuronu bude pro perceptron:

$$y = \begin{cases} 0 & \text{pro } S < 0 \\ 1 & \text{pro } S \geq 0 \end{cases}, \quad (5.4)$$

a vnitřní potenciál (původní rovnice (5.2)) bude nově vyjádřen jako:

$$S = \sum_{i=1}^n w_i x_i + \theta. \quad (5.5)$$

Lze zobecnit výše uvedený matematický zápis výstupu perceptronu na

$$y = f_N(S), \quad (5.6)$$

kde f_N je nelineární aktivační funkce neuronu. Do množiny těchto funkcí patří kromě funkce (5.4) mnoho různých funkcí, které budou podrobněji popsány v dalších kapitolách.

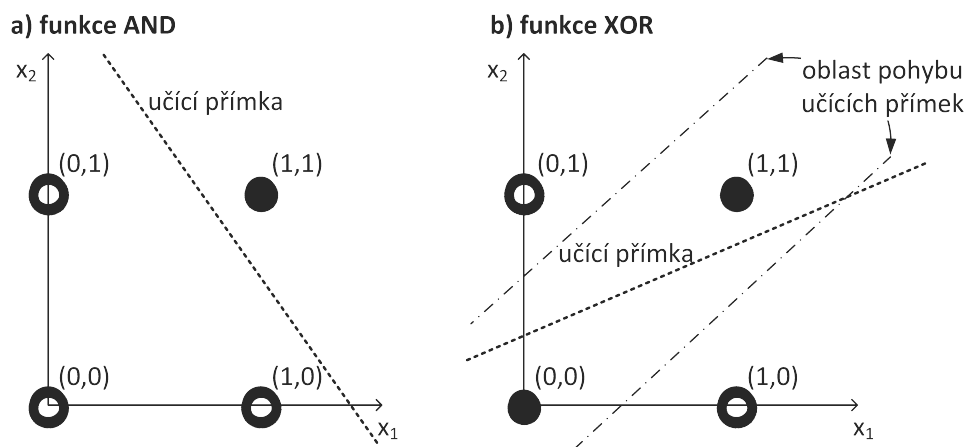
K perceptronu v 60. letech 20. století vznikaly učící metody, které jsou obvykle zaměřeny na strojové rozpoznávání (postup učení a jeho zhodnocení je uveden například v [76] a [77]). Tyto metody byly zaměřeny pouze na samotné perceptrony, případně více peceptronů v jedné vrstvě. Tyto první sítě tedy mohly mít více výstupů než jeden. Hlavním důvodem, proč nebylo možné budovat sítě vícevrstvé, byla - v době vzniku - příliš velká výpočetní náročnost těchto učících algoritmů. Zástupcem neuronových sítí s učícími algoritmy jsou například sítě ADALINE (ADaptive LInear NEuron) a v 80. letech MADALINE (Many ADALINE) popsané například v pracích B. Widrowa [78] a [79].

V 70. a 80. letech dochází k velkému úpadku vědeckého zájmu o celý obor umělých neuronových sítí. Tento pokles zájmu je často připisován knize Minského a Paperta - Perceptrons, z roku 1969 ([80]). V této knize, kromě formalizace dosavadních znalostí o neuronech a jejich strukturách, autoři došli k závěru, že neuronové sítě nejsou tak univerzální jak se do této doby přepokládalo a že nejsou schopny řešit určitou velkou množinu problémů. Později je tento závěr označován jako Minsky-Papertův omyl. Minsky a Papert se v knize mimo jiné zabývali způsobem, jakým se dají použít perceptrony ke klasifikaci - tj. rozpoznání zástupců dvou množin v jednom stavovém prostoru. Nejjednodušším případem je neuron se dvěma binárními vstupy (to znamená, že celý stavový prostor tvoří nejvýše 4 jedinci - (0,0), (0,1), (1,0) a (1,1)) a každý tento jedinec může patřit do jedné ze dvou množin (ty představuje výstup neuronu 1 nebo 0 resp. -1). Průběh učení neuronu lze geometricky interpretovat jako přímku, která se v prostoru posouvá a otáčí. To je zajištěno úpravami vah pomocí učícího algoritmu (v tomto případě může být použit libovolný učící algoritmus). Na začátku učícího procesu je učící přímka umístěna náhodně podle toho, na jaké hodnoty jsou inicializovány váhy w_1 , w_2 a prahová hodnota θ na vstupu neuronu. Učící algoritmus postupně mění váhy tak, aby učící přímka, kterou lze definovat rovnicí:

$$x_2 = \frac{-w_1}{w_2}x_1 + \frac{\theta}{w_2}, \quad (5.7)$$

rozdělila stavový prostor na dvě požadované množiny, jak je zobrazeno na Obr. 5.2a. Zásadní problém je však s funkcí XOR, kdy není možné nastavit přímku v prostoru tak, aby byl prostor správně rozdělen na dvě množiny (to je zobrazeno na Obr. 5.2b). Lze tedy říci, že jeden neuron není schopen správně klasifikovat lineárně neseparovatelné množiny. Toho by v daném případě bylo možné dosáhnout přidáním dalších neuronů, ale pak by analogicky bylo možno najít další případ (už s více vstupy), kdy nebude ani tento počet neuronů stačit. V důsledku toho bylo z práce Minského a Paperta nevhodně vyvozeno, že neuronové sítě nebudou schopny správně fungovat pro všechny případy klasifikace. Jak bylo později v roce 1987 dokázáno v práci [81], tento problém lze vyřešit pomocí vícevrstvé neuronové sítě. Masovější návrat ke studiu neuronových sítí se tedy obnovil až v druhé polovině 80. let. Od té doby probíhá studium neuronových sítí jak z pohledu jejich architektury a aplikací, tak i jejich HW implementací v podobě neuročipů a SW

implementací jako je Neural Network Toolbox pro Matlab, NeuroSolutions nebo Stuttgart Neural Network Simulator.



Obr. 5.2: Grafická reprezentace klasifikace funkce AND a problému s funkcí XOR

V následující kapitole budou popsány základní neuronové sítě včetně všech sítí použitých v této práci a principy jejich učení a vybavování. Budou rovněž popsány některé z podstatných aplikací těchto sítí.

5.2 Druhy neuronových sítí

Neuronové sítě je možné rozdělit z pohledu architektury na jednovrstvé, vícevrstvé a rekurentní. Jak již z názvu vyplývá, tyto sítě se liší počtem vrstev neuronů, kterými prochází vstupní informace, dokud neuronová síť neposkytne požadovaný výstup. Mnoho z těchto sítí má těžiště svého využití především v oblasti strojového rozpoznávání obrazu, zvuku apod.. Nehledě na tuto skutečnost jsou v následujícím výčtu uvedeny, ačkoli se v této práci nemusí ve své čisté podobě vyskytovat, protože jejich architektura byla inspirací pro další sítě a především obsahují důležité myšlenky využitě v učících algoritmech dalších sítí.

Mezi jednovrstvé sítě patří:

- Perceptron
- Adaline
- Hopfieldova síť
- Kohonenova síť (Self Organizing Map - SOM)

Nejpoužívanějšími vícevrstvěmi sítěmi jsou:

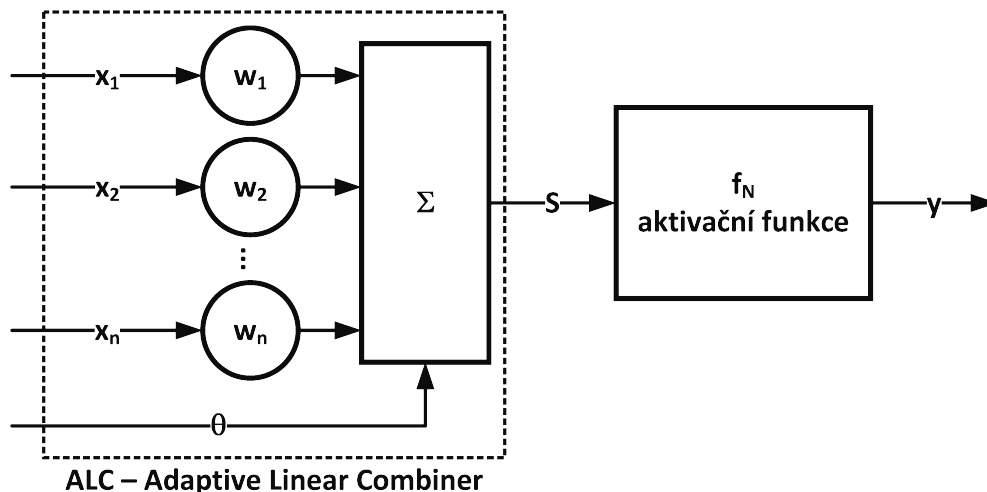
- Multi-layer (ML) Perceptron
- RBF síť

- Síť ART (Adaptive Resonance Theory)
- Cognitron/Neocognitron

Rekurentní sítě jsou obecně sítě typu ML, ve kterých se v propojení jednotlivých vrstev objevují propoje přes zpožďovací prvek do některé z předchozích vrstev.

5.2.1 Perceptron

Perceptron je základní jednovrstvá síť, která používá Rosenblatovské perceptrony, které byly částečně popsány už v předchozím textu v sekci 5.1. Základem je přenosová funkce perceptronu uvedená v rovnicích (5.5) a (5.6). V tomto zápisu je vidět, že jediný stupeň volnosti definice perceptronu je tvořen aktivační funkcí f_N , která je ale zásadní pro správnou činnost neuronové sítě. Graficky je perceptron znázorněn na obr. 5.3



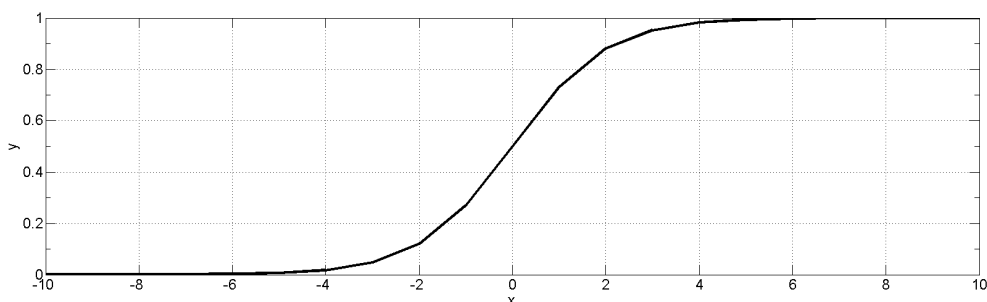
Obr. 5.3: Schéma funkce jednoho perceptronu

5.2.1.1 Aktivační funkce neuronu

Vstupem aktivační funkce neuronu je vždy vnitřní potenciál S , tedy výstup bloku ALC (Adaptive Linear Combiner), který zajišťuje výpočet lineární kombinace vstupů a příslušných vah. Aktivační funkce je obvykle nelineární funkce, která zajišťuje, aby se výstup neuronu držel v požadovaném intervalu hodnot. Tento interval může být unipolární - nesymetrický, např. $(0, 1)$ - nebo bipolární, např. $(-1, 1)$. Nejběžnější aktivační funkcí je logaritmická sigmoida (v MATLABu ji představuje funkce $\text{logsig}()$), která je spojitě diferencovatelná v celém definičním oboru a je definována vztahem:

$$f_N(x) = y = \frac{1}{1 + e^{-\alpha x}}, \quad (5.8)$$

kde parametr α ovlivňuje přimknutí křivky ke svislé ose - pro menší α se křivka napřimuje a strmost klesá. Tento parametr také v některých případech ovlivňuje činnost neuronové sítě v okamžiku, kdy váhy neuronu dosahují velkých hodnot. V takovém případě je vhodné zmenšit tento parametr, aby se sigmoida protáhla přes větší interval a umožnila tak urychlit učící proces (ten je závislý na derivaci aktivační funkce, což bude podrobněji popsáno v kapitole zabývající se algoritmem Back-propagation). Pro aktivační funkci (5.8) platí vlastnosti: $\{x \rightarrow -\infty\} \iff \{y \rightarrow 0\}$; $\{x = 0\} \iff \{y = 0.5\}$; $\{x \rightarrow \infty\} \iff \{y \rightarrow 1\}$. Průběh funkce je zobrazen na obr. 5.4.



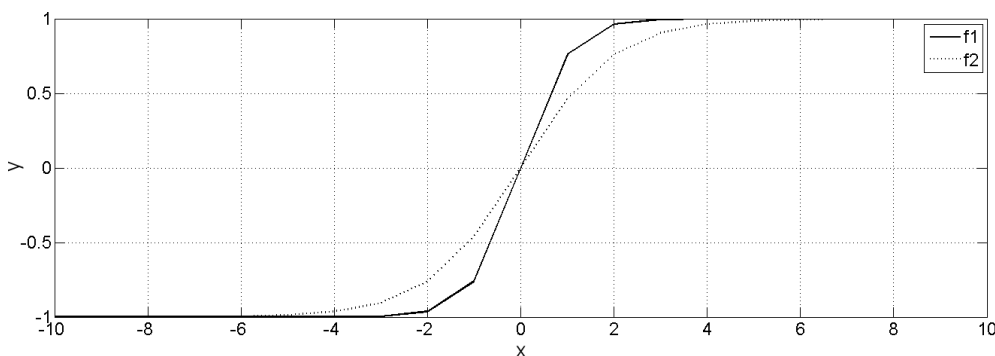
Obr. 5.4: Průběh funkce unipolární sigmoidy podle (5.8) s parametrem sigmoidy $\alpha = 1$

Dalšími často používanými bipolárními aktivačními funkcemi jsou například hyperbolická tangenta:

$$f_N(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5.9)$$

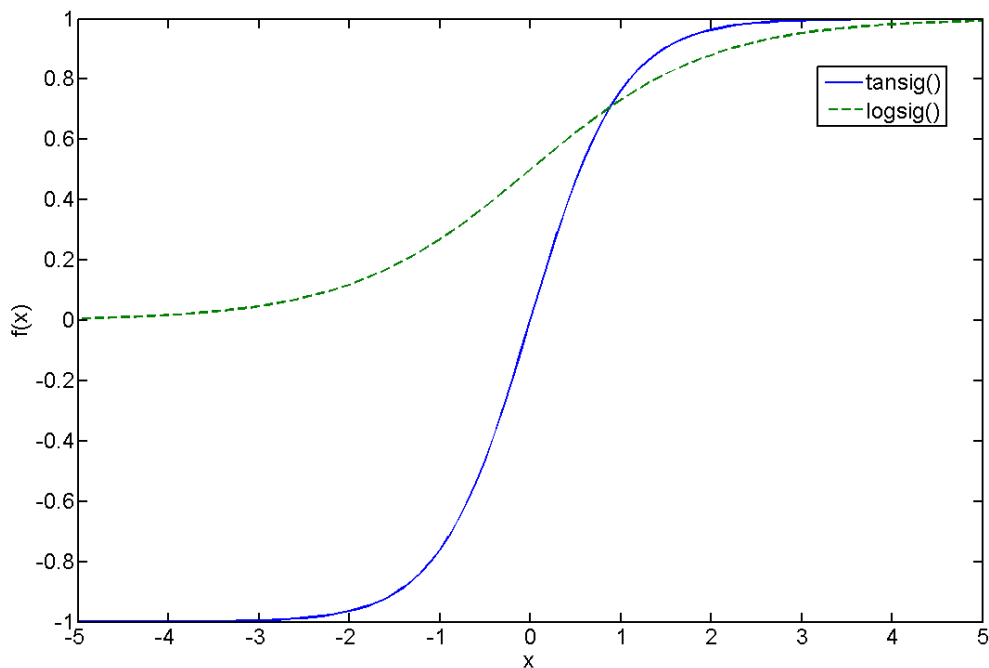
případně i upravená původní unipolární funkce (5.8) tak, aby byl její obor hodnot $(-1, 1)$:

$$f_N(x) = y = \frac{2}{1 + e^{-x}} - 1. \quad (5.10)$$

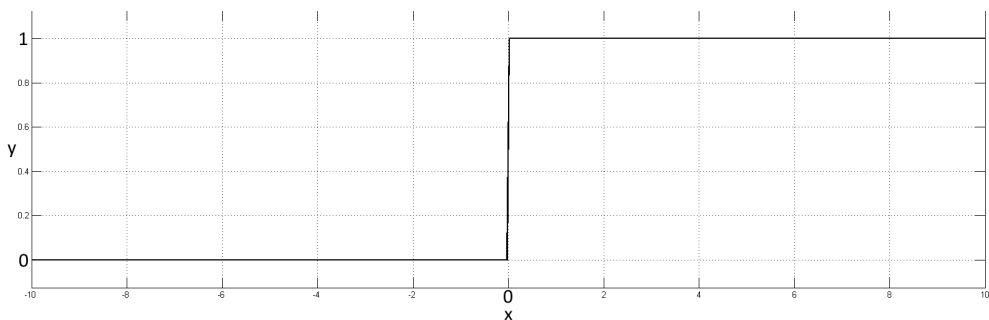


Obr. 5.5: Průběh funkce bipolárních sigmoid: $f1 = \tanh(x)$, $f2 =$ funkce (5.10)

V systému Matlab a Neural Network Toolboxu, který bude využíván dále v této práci, jsou typickými aktivačními funkcemi `logsig()` a `tansig()`, jejich srovnání je uvedeno na obr. 5.6.

Obr. 5.6: Srovnání funkcí $tansig()$ a $logsig()$

V neuronových sítích, u kterých se požadují ostré výstupy - typicky při strojovém rozpoznávání a ve výstupní vrstvě se velmi často využívají aktivační funkce, které nejsou spojitě diferencovatelné, ale lépe vystihují podstatu problému. Jsou to například Heavisideova funkce (zobrazena na obr. 5.7) a jiné obdobné skokové funkce. V některých implementacích neuronu může být tato funkce posunuta ve směru osy x - tento posun pak představuje prahovou hodnotu neuronu θ .



Obr. 5.7: Heavisideova funkce

5.2.1.2 Aplikační možnosti jednovrstvé sítě perceptronového typu

Tyto sítě jsou schopny řešit lineárně separovatelné problémy, jak již bylo popsáno v kapitole 5.1 u popisu Minsky-Papertova omylu. Je tedy třeba určit, jak velká je oblast lineárně separovatelných problémů. Pro neuron, který má n vstupů, může mít celý stavový prostor velikost 2^{2^n} (u každé kombinace vstupů má určit, zda patří do jedné nebo druhé množiny). Kvantifikací lineárně separovatelných

problémů se zabýval R. O. Winder ve svém příspěvku [82] a došel k závěru, že pro počet lineárně separovatelných problémů R_n platí vztah:

$$R_n \leq 2 \sum_{i=0}^n \binom{2^n - 1}{i}, \quad (5.11)$$

pro který platí přibližně

$$R_n \leq 2 \binom{2^n}{n}. \quad (5.12)$$

V knize [83] je pak ještě uváděna hrubá aproximace počtu lineárně separovatelných problémů $R_n < x^{1/3}$ pro $n > 7$.

5.2.1.3 Učení jednovrstvé sítě perceptronovského typu

Pro perceptronovské sítě je typické učení s učitelem (supervised learning). Jeho myšlenka spočívá ve srovnání aktuálního výstupu s požadovaným výstupem a nastavení jednotlivých vah tak, aby se chyba snižovala. Algoritmus pro učení perceptronu se skokovou aktivační funkcí má následující průběh:

1. Inicializace vektoru vah \mathbf{w} perceptronu náhodnými hodnotami
2. Výpočet výstupu neuronu a jeho porovnání s požadovanou hodnotou
3. Je-li aktuální hodnota shodná s požadovanou, váhy se nemění a pokračuje se učením dalšího vstupu.
4. Pokud se liší aktuální a požadovaný výstup, změní se váhy na vstupech. Přístupy k této změně mohou být například:
 - Aplikují se pevné úbytky (pokud rozdíl aktuální a požadované hodnoty je kladný) resp. přírůstky (pokud je rozdíl záporný).
 - Velikost změny se mění v závislosti na velikosti chyby. Tím je možno zrychlit konvergenci k řešení, ale zároveň se může snížit stabilita algoritmu. Velmi často bývá zaveden učící koeficient z intervalu $(0, 1)$, kterým se vypočtená chyba výstupu násobí.
 - Lze rovněž kombinovat oba předchozí přístupy.
5. Pokračuje se v učení dalšího vstupu

Výše uvedený algoritmus může být pro neuron s menším počtem vstupů rovněž velmi názorně graficky zobrazen, jak je uvedeno v [84]. Pro lepší porovnání myšlenky tohoto algoritmu s předchozím popisem neuronu a Minsky-Papertova omylu z kapitoly 5.1 bude uvažován neuron se dvěma vstupy. Jednotlivé prvky stavového prostoru jsou v rovině a učení neuronu představuje přímkou v této rovině. Algoritmus je ale obecný pro libovolný počet vstupů - např. pro tři vstupy

by stavový prostor představoval trojrozměrný prostor a učení by pohybovalo rovinou, která by množiny separovala. Rovnici přímky separující dvě množiny lze napsat ve tvaru:

$$w_1x_1 + w_2x_2 + w_3 = 0, \quad (5.13)$$

kde x_i jsou vstupní hodnoty perceptronu, w_1 a w_2 jsou váhy a w_3 může představovat prahovou hodnotu. Vektor $\mathbf{w} = (w_1, w_2)$ je tedy normálový vektor separační přímky. Všechny vstupní stavy neuronu patří do jedné ze dvou lineárně separovatelných množin A nebo B . Cílem učícího algoritmu je tedy najít takový vektor \mathbf{w} , pro který bude platit, že:

$$\begin{aligned} \mathbf{w} \cdot \mathbf{x}_i &> 0 \text{ pro vstupy z množiny } A \\ \mathbf{w} \cdot \mathbf{x}_i &< 0 \text{ pro vstupy z množiny } B, \end{aligned} \quad (5.14)$$

kde vstupní vektor $\mathbf{x}_i = (x_1, x_2)$. Platí tedy, že směrové vektory jednotlivých vstupních stavů neuronu \mathbf{x}_i jsou ve stejném směru jako normálový vektor příslušné učící přímky. Postup algoritmu je následující:

1. Učící algoritmus začíná s náhodně nastavenými hodnotami vektoru \mathbf{w} .
2. Pro učení jsou v náhodném pořadí vybírány vektory \mathbf{x}_i z celé množiny $A \cup B$.
3. Je vypočten skalární součin $\mathbf{w} \cdot \mathbf{x}_i$. Zde je třeba rozlišit, patří-li vektor do množiny A nebo B .

$\mathbf{x}_i \in A$: Pokud je skalární součin kladný, není prováděna změna vah a pokračuje se krokem 2.

Jinak: vektor nových vah je vypočten jako: $\mathbf{w}_{t+1} = \mathbf{w}_t + \mathbf{x}_i$.

$\mathbf{x}_i \in B$: Pokud je skalární součin záporný, není prováděna změna vah a pokračuje se krokem 2.

Jinak: vektor nových vah je vypočten jako: $\mathbf{w}_{t+1} = \mathbf{w}_t - \mathbf{x}_i$.

4. Pokračuje se krokem 2 dokud existuje \mathbf{x}_i , pro které se mění vektor vah (tj. dokud existuje vstup, který není správně klasifikován).

V průběhu algoritmu tedy dochází k rotaci vektoru \mathbf{w} , ale vzhledem k tomu, že vektory nejsou normalizované, dochází vždy k rotaci o jiný úhel. Tento způsob učení má tu vlastnost, že velikost vektoru \mathbf{w} průběžně roste a korekce uvnitř učícího algoritmu jsou menší. To značí i to, že rychlost učení (learning rate) také klesá až k nule. To je typický rys i mnoha jiných učících algoritmů.

Ohledně rychlosti konvergence učících algoritmů bylo vypracováno mnoho prací, ale nelze najít univerzální řešení pro všechny problémy. Například v práci [85] byly porovnány různé učící algoritmy s lineárním programováním a bylo dokázáno, že počet učících epoch roste v závislosti na počtu vstupů v řádu $O(n^3 \log(n))$, ale pro nejhorší případ $O(2^n)$. V praktických testech učících algoritmů ale bylo dosahováno výrazně lepších výsledků. Na druhou stranu bylo rovněž doporučeno,

aby jako horní hranice počtu učících cyklů v běžných případech bylo považováno 10^7 - pro tento počet je velmi pravděpodobné, že velikosti vah mohou dosáhnout velikosti požadované změny vah a algoritmus nemusí konvergovat k řešení. V práci [86] bylo rovněž dokázáno, že složitost učení perceptronové sítě může být i polynomiální a závislá na požadované přesnosti.

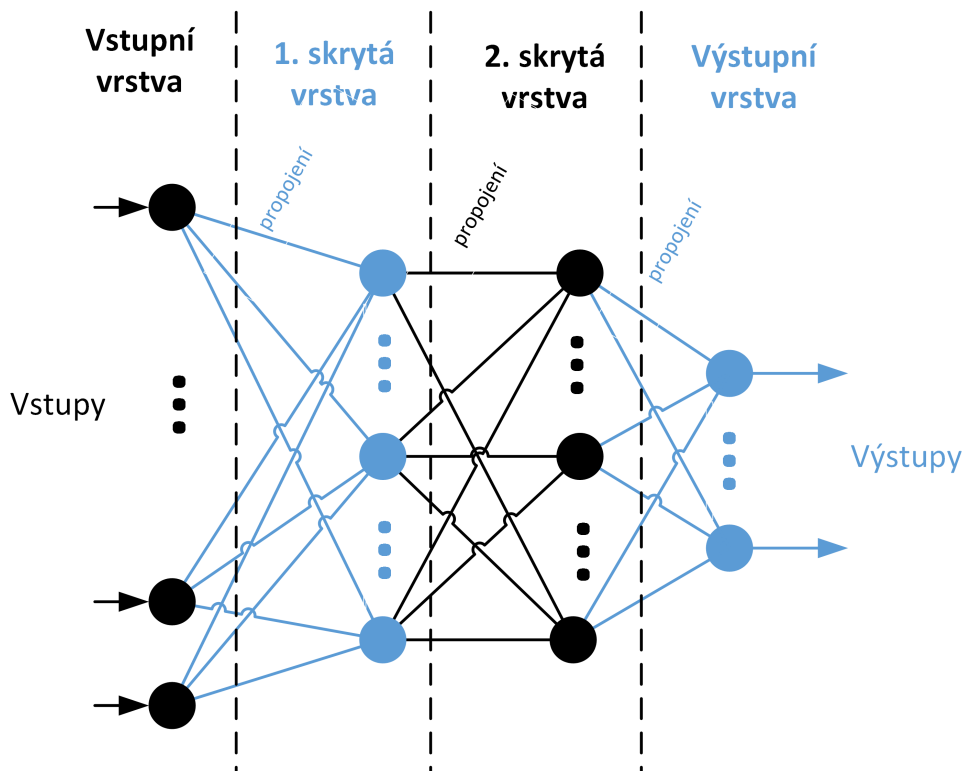
5.2.2 Multi-layer Perceptron

Jak již bylo zmíněno v předchozích odstavcích, jednovrstvá síť je schopna vyřešit klasifikační problém, kde se prvky jedné množiny nacházejí nejhůře v otevřené konvexní oblasti. V 80. letech v práci [87] bylo ukázáno, že dvouvrstvá síť dokáže vyřešit problém funkce XOR. Lze zobecnit, že tato síť dokáže klasifikovat problémy s body v nekonvexní oblasti - myšlenka je taková, že neurony v jedné vrstvě představují konvexní oblasti problému a další vrstva vytváří průnik oblastí v předchozí vrstvě. Rozšířením tohoto postupu je, že neuronová síť s třemi vrstvami už dokáže klasifikovat obecné oblasti (průnik nekonvexních oblastí). Třívrstvá neuronová síť, schopná vyhodnotit obecný klasifikační problém, je zobrazena na obr. 5.8. Je třeba zdůraznit, že u vícevrstevných neuronových sítí se předpokládá, že vrstvy jsou mezi sebou propojeny v pořadí, v propojení neuronů neexistují cykly ani zpětné propoje mezi vrstvami, proto se tyto sítě také často pojmenovávají jako dopředné neuronové sítě (Feed-Forward Artificial Neural Networks).

První vrstva MLP se nazývá vstupní. Neurony v této vrstvě nemají žádné předchůdce a je na ně přiváděn vstupní signál. Jde vlastně jen o velmi zjednodušené perceptrony, které mají za úkol rozvést vstupní signál na více perceptronů v další vrstvě bez změny hodnoty.

Následují skryté vrstvy. Jejich počet není prakticky omezen, ale pro obecný klasifikátor postačují dvě skryté vrstvy, jak bylo popsáno v předchozích odstavcích. V některých případech mohou další skryté vrstvy vylepšit rychlost nebo kvalitu učení neuronové sítě. Při vyšším počtu skrytých vrstev ale obecně roste náročnost učení. V této vrstvě jsou už plnohodnotné neurony se všemi vahami, prahy a aktivačními funkcemi.

Poslední vrstvou dopředných sítí jsou výstupní vrstvy. Tyto neurony mají za úkol poskytnout požadovanou výstupní informaci. Podle druhu zamýšleného využití sítě mívají neurony v této vrstvě buď spojitou (sigmoidní) aktivační funkci nebo skokovou aktivační funkci. V případě, že síť je určena ke klasifikaci, bývají výstupní neurony spíše dvoustavové a tudíž mají skokovou aktivační funkci (například Heavisideova na obr. 5.7). Pokud má síť fungovat jako aproximátor nějaké reálné funkce, bývají výstupní neurony se spojitým výstupem, tudíž aktivační funkce by měla být spojitá, což splňuje jak sigmoida (obr. 5.4 a 5.5), tak i třeba lineární funkce. Tyto neurony tedy udávají i požadovaný rozsah výstupních hodnot.



Obr. 5.8: Vícevrstvá síť perceptronovského typu (Multi-Layer Perceptron - MLP)

5.2.2.1 Učení MLP sítě

Problémem MLP sítě byla v době vzniku nedostupnost učícího algoritmu. Existoval algoritmus pro vícevrstvé síť Madaline, který ale nebyl dostatečně výkonný pro učení obecné vícevrstvé sítě. V práci [87] je kromě vyřešení problému XOR popsáno zobecněné delta-pravidlo (Generalized Delta-Rule), které je možné použít pro učení vícevrstvé sítě. Tento postup je inspirován původním učícím algoritmem delta-pravidlem publikovaným v [88]. Je to jedna z gradientních metod, která je zvláštním případem univerzálního algoritmu Back-Propagation (algoritmus zpětného šíření), který je asi nejčastěji používán v dopředných vícevrstevných neuronových sítích. Aplikaci zobecněného delta-pravidla lze rozdělit do dvou fází:

1. Přivedení vstupních hodnot a výpočet odezvy na tento vstup. Je vypočten rozdíl δ požadovaného a aktuálního výstupu.
2. Zpětný průchod celou sítí, během kterého je v předchozí fázi vypočtená chybová hodnota předána každému neuronu a jsou podle něho upravovány váhy. Způsob úpravy vah a jeho odvození je uvedeno v [87]. Změny vah mohou být rekurzivně spočteny pro všechny předcházející vrstvy neuronové sítě v okamžiku, kdy jsou známy chybové hodnoty vrstvy následující.

Proces zpětného průchodu neuronovou sítí má přibližně stejnou výpočetní náročnost jako dopředný průchod neuronovou sítí.

5.2.2.2 Algoritmus back-propagation (BP)

Tento algoritmus byl poprvé publikován v roce 1986 v práci [89] a jedná se o iterační učení s učitelem. Tento algoritmus v 80. letech nastartoval nový rozvoj v oblasti neuronových sítí po odklonu od neuronových sítí, který nastal po vydání práce Minskeho a Paperta. Průběh učení je obdobný jako u učení jednovrstvé sítě jen s tím rozdílem, že modifikace vah je složitější. Cílem učení je minimalizace globální chyby sítě, která je definována jako střední kvadratická odchylka (RMS - Root Mean Squared) odezev na všechny vstupy z trénovací množiny. Tato odchylka je definována jako:

$$RMS = \sum_m \sqrt{\sum_n (y_{mn} - d_{mn})^2}, \quad (5.15)$$

kde m je velikost trénovací množiny, n je počet neuronů ve výstupní vrstvě, y_{mn} je odezva na zadaný vstup na n -tém výstupním neuronu pro m -tý vzor z trénovací množiny a d_{mn} je požadovaná odezva na n -tém neuronu pro m -tý vzor z trénovací množiny. Trénování probíhá v jednotlivých iteracích a průběh jednoho učení všech vzorků trénovací množiny se nazývá **epocha**. Trénovací množina nemusí být vždy tvořena všemi učícími vzory, které má síť klasifikovat, ale v mnoha případech je možné vybrat jen určitou část z učících vzorů. Konstrukcí trénovací množiny se budeme zabývat později. Samotný výběr vzorů z trénovací množiny nemusí být sekvenční. Tento výběr může být náhodný, dokonce některé vzory mohou být na vstup přivedeny v rámci jedné epochy několikrát. Důležité je však to, že během jedné epochy je každý vzor učen alespoň jednou.

Délku učení v praktických případech nelze určit dopředu - obvykle se učení zastavuje po splnění předem určené podmínky, kterou může být konkrétní hodnota globální chyby (RMS), průměrná chyba RMS vztažená na jeden vzorek trénovací množiny nebo trénovací čas, případně počet epoch. Průběh učení se obvykle sleduje na průběhu globální chyby v závislosti na počtu epoch, případně i na průběhu chyby na testovací množině. Testovací množina není pro učení sítě nezbytná a bývá obvykle konstruována obdobným způsobem jako trénovací, ale neměla by s ní být shodná. Užívá se po určitém počtu epoch k ověření kvality učení. Pokud klesá globální chyba u trénovací i testovací množiny, lze říci, že učení probíhá správně. Pokud se průběh chyby obou množin výrazně liší, lze podle toho detekovat buď chybu v návrhu sítě, chybně sestavenou trénovací množinu nebo tzv. přeučení sítě (overfitting). To nastává například v případě, kdy je trénovací množina příliš malá nebo neobsahuje správné množství vzorků a naučená síť tak nemá schopnost generalizace.

Vzhledem k důležitosti algoritmu Back-propagation se v tuto chvíli zaměříme na podrobnější popis tohoto algoritmu a vysvětlení způsobu, jakým se upravují jednotlivé váhy v dopředné síti. Budeme předpokládat, že existuje energetická

(chybová) funkce neuronové sítě, která vychází z metody nejmenších čtverců:

$$E = \frac{1}{2} \sum_m \sum_n (y_{mn} - d_{mn})^2, \quad (5.16)$$

kterou se učení snaží minimalizovat za pomoci sestupu ve směru nejmenšího gradientu. Tuto funkci si můžeme představit jako plochu v hyperprostoru. Obecně je velmi složitá a závislá na vahách a prazích jednotlivých neuronů a vstupních a výstupních hodnotách neuronové sítě. Lze vyjádřit chybovou funkci E_k pro jeden výstup k :

$$E_k = \frac{1}{2} \sum_k (y_k - d_k)^2. \quad (5.17)$$

Učící algoritmus je tedy optimalizační úloha, která má nalézt správné váhy a prahy jednotlivých neuronů. Každý jednotlivý neuron může mít jinou strukturu, ale v následujícím textu budeme předpokládat, že splňuje rovnice (5.5) a (5.6) a aktivační funkcí je logaritmická sigmoida podle rovnice (5.8). Pro přehlednost si zavedeme jednotné značení jednotlivých prvků neuronové sítě. Například bude platit

$$w_{ij}^{[vrstva]} \quad (5.18)$$

pro váhový koeficient j -tého vstupu v i -tém neuronu ve vrstvě $[vrstva]$. Obdobně to bude platit pro výstup neuronu $y_i^{[vrstva]}$. Pro přehlednost bude aktivační funkce ((5.6)) přeznačena na $F(\Psi)$, kde Ψ je vnitřní potenciál neuronu ((5.5)).

Pro výpočet změn vah uvnitř neuronové sítě potřebujeme znát lokální gradient energetické funkce E , protože pro změny všech vah platí:

$$\Delta w = -\eta \nabla E. \quad (5.19)$$

Hodnota η je učící parametr, který určuje velikost kroku po energetické funkci. Tímto parametrem lze výrazně ovlivnit průběh učení a je zřejmé, že tento parametr musí být ostře větší než 0. Vhodnou volbou lze překonat mělká lokální minima, ale na druhou stranu velká hodnota může způsobit oscilace sítě a pomalou konvergenci sítě. Lokální gradient lze rozepsat pro výstupní vrstvu:

$$\frac{\partial E}{\partial w_{ij}^{[vystupni]}} = \frac{\partial E}{\partial y_i^{[vystupni]}} \frac{\partial y_i^{[vystupni]}}{\Psi_i^{[vystupni]}} \frac{\Psi_i^{[vystupni]}}{\partial w_{ij}^{[vystupni]}} \quad (5.20)$$

Z derivace energetické funkce (5.17) dostaneme:

$$\frac{\partial E}{\partial y_i^{[vystupni]}} = (y_i^{[vystupni]} - d_i). \quad (5.21)$$

Dále je třeba znát derivaci aktivační funkce (sigmoidy):

$$\frac{\partial y_i^{[vystupni]}}{\Psi_i^{[vystupni]}} = \alpha y_i^{[vystupni]} (1 - y_i^{[vystupni]}). \quad (5.22)$$

Poslední faktor rovnice (5.20) vychází z derivace vnitřního potenciálu a z toho, že do poslední vrstvy jako vstup přichází výstup neuronu z předchozí vrstvy:

$$\frac{\Psi_i^{[vystupni]}}{\partial w_{ij}^{[vystupni]}} = y_j^{[skryta2]}. \quad (5.23)$$

Lokální gradient (5.20) lze tedy rozepsat jako:

$$\frac{\partial E}{\partial w_{ij}^{[vystupni]}} = (y_i^{[vystupni]} - d_i) \alpha y_i^{[vystupni]} (1 - y_i^{[vystupni]}) y_j^{[skryta2]} = \delta_i^{[vystupni]} y_j^{[skryta2]}. \quad (5.24)$$

Lze již tedy napsat změnu vah ve výstupní vrstvě:

$$\Delta w_{ij}^{[vystupni]} = -\eta \delta_i^{[vystupni]} y_j^{[skryta2]}. \quad (5.25)$$

Výpočet změn vah ve skrytých vrstvách naráží na problém, že neznáme hodnoty, kterých mají výstupy neuronů v této vrstvě nabývat. Chybu neuronů ve skryté vrstvě lze určit ze součtu chyb výstupní vrstvy (všech neuronů) pokud ji vynásobíme příslušnými vahami. Tento přístup lze zobecnit pro všechny další vrstvy. Zde tedy vidíme, že se vypočtená chyba šíří zpět přes všechny vrstvy - odtud pochází název Back-propagation. Pro výpočet koeficientu změny vah $\delta_i^{[skryta2]}$ budeme tedy postupovat stejně jako v prvním kroku a spočteme gradient:

$$\frac{\partial E}{\partial w_{ij}^{[skryta2]}} = \frac{\partial E}{\partial y_k^{[vystupni]}} \frac{\partial y_k^{[vystupni]}}{\Psi_k^{[vystupni]}} \frac{\Psi_k^{[vystupni]}}{\partial y_i^{[skryta2]}} \frac{\partial y_i^{[skryta2]}}{\Psi_i^{[skryta2]}} \frac{\Psi_i^{[skryta2]}}{\partial w_{ij}^{[skryta2]}} = (5.26)$$

$$= \delta_k^{[vystupni]} \delta_{ki}^{[vystupni]} \frac{\partial y_i^{[skryta2]}}{\Psi_i^{[skryta2]}} \frac{\Psi_i^{[skryta2]}}{\partial w_{ij}^{[skryta2]}} = \delta_i^{[skryta2]}, \quad (5.27)$$

kde k je index neuronu v následné vrstvě (odkud se šíří již vypočtená chyba). Po přičtení všech příspěvků chyby z výstupní vrstvy platí

$$\delta_i^{[skryta2]} = y_i^{[skryta2]} (1 - y_i^{[skryta2]}) \sum_k w_{ki}^{[vystupni]} \delta_k^{[vystupni]}. \quad (5.28)$$

Pro adaptaci vah tedy bude platit:

$$\Delta w_i^{[skryta2]} j = -\eta \delta_i^{[skryta2]} y_j^{[skryta1]}. \quad (5.29)$$

V předchozím textu byl ukázán průběh učícího algoritmu a zároveň byly zmíněny parametry, které jsou důležité pro samotné učení. Existuje celá řada postupů jak je možné tento algoritmus zrychlit, v některých případech zlepšit stabilitu a podobně. Jedním z takových přístupů je zavedení setrvačnosti ve změnách vah. Tento přístup dokáže zlepšit stabilitu průchodu energetickou hyperplochou k minimu tím, že nové váhy budou vyhodnoceny podle rovnice:

$$\Delta \mathbf{w}(T+1) = -\eta \delta y + \mu \Delta \mathbf{w}(T). \quad (5.30)$$

Druhý člen této rovnice zahrnuje parametr servačnosti μ , který by se měl pohybovat v intervalu $< 0, 1 >$. Je tím zajištěno částečné udržení směru pohybu z předchozích učících epoch a pohyb po hyperploše je tak plynulejší a má větší šanci překonávat mělká lokální minima, ve kterých by jinak učení mohlo uváznout. Rovněž průchod k minimu je rychlejší než při použití standardního BP algoritmu.

Další možností jak ovlivnit učení sítě:

- Změna parametru α , pokud je aktivační funkce ve tvaru sigmoidy. Tento postup se často používá pro paralyzované sítě, u kterých učení konverguje příliš pomalu.
- V případě uváznutí v lokálním minimu (detekováno průběhem učení) je možnost použít šokovou změnu vah o malou náhodnou hodnotu.
- Jinou alternativou řešení uváznutí v lokálním minimu může být simulované ochlazování - po uváznutí je povoleno přijímat i horší řešení, dokud se učení nedostane z minima. Pak znovu nastoupí BP algoritmus.

Významnými modifikacemi BP algoritmu jsou například:

- Fast Backpropagation - publikováno v [90], kdy se před aktualizací vah přidává chyba příslušné vrstvy
- Quickprop - jeden z nejefektivnějších algoritmů pro učení sítě, kdy jsou pro změnu vah využívány i derivace vyšších řádů, čímž je v plochých částech aktivační funkce zrychlen pohyb po energetické hyperploše. Tento přístup byl publikován v [91].

5.2.3 Kohonenova síť - SOM

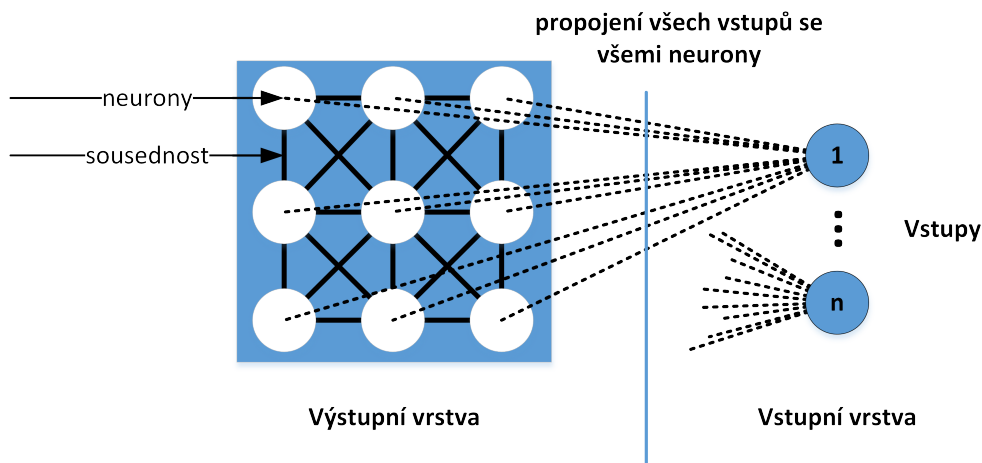
Kohonenova síť má zcela jinou architekturu než sítě uvedené v předcházejících kapitolách a její použití je zaměřeno na řešení jiných druhů problémů. Těmi jsou především klasifikace a analýza dat. Síť je učena bez učitele a patří do skupiny samoorganizujících neuronových sítí. Lze tedy zjednodušeně říci, že sama dokáže rozlišovat objekty s podobnými vlastnostmi. Samoorganizací biologických neuronů ve vztahu ke vnímání vnějších podnětů se v 70. letech zabývali především Christoph von der Malsburg ve spolupráci s Davidem J. Willshawem v celé řadě publikací - například [92], [93], [94] a [95]. Lze říci, že jejich práce je velkou motivací pro sítě, které od 80. let vyvíjí Teuvo Kohonen a podle něj jsou i tyto sítě pojmenovány. Jedná se o Kohonenovy samoorganizující mapy (SOM). Tyto sítě jsou rovněž vhodné pro vizualizaci vícerozměrných množin dat. Myšlenka samoorganizace u Kohonenových sítí je podobná lidskému mozku (to je pouze biologická domněnka) - podnětová centra jsou v mozku rozmístěna v celém objemu a existují 2 předpoklady:

- Na konkrétní vstup reagují neurony v určitých konkrétních oblastech.

- Na opakovaný nebo velmi podobný vstup reagují stejné neurony a zároveň se stávají citlivějšími na tyto podněty.

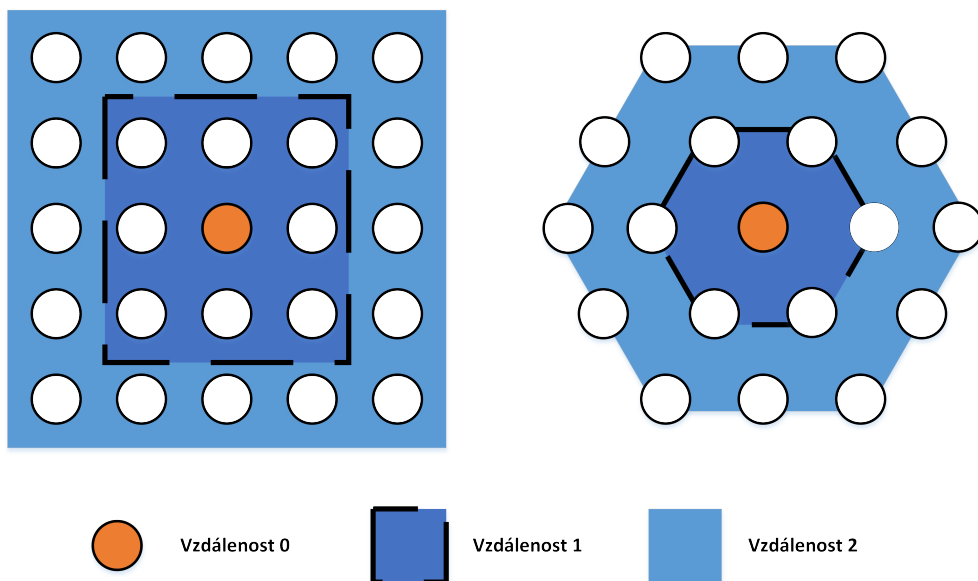
Tento princip je také využit v sítích SOM, což bude rozebráno dále v textu. Architektura SOM sítě je jednovrstvá, ale neobsahuje neurony stejného typu jako u dopředných neuronových sítí. Vstupní n -rozměrný vektor přichází do vstupní vrstvy, která není plnohodnotná - zajišťuje pouze rozvětvení dat vstupního signálu, aby se každá složka vstupu dostala ke každému neuronu z výstupní (Kohonenovy) vrstvy. Schéma této sítě je na obr. 5.9. Nejdůležitější je tedy u SOM výstupní vrstva, ve které jsou neurony uspořádány v sousednosti. Toto uspořádání neuronů je relativně volné a definuje topologii, která bude využívána při učení sítě. V topologii musí být možné definovat vzdálenosti jednotlivých neuronů. Na obr. 5.10 jsou zobrazeny jedny z nejčastěji používaných definic topologie pro dvourozměrnou výstupní vrstvu. Okolí se liší ve tvaru a tedy i v počtu sousedních neuronů. Výhodou víceúhelníkového okolí (např. na obr. 5.10b) je větší podobnost s kružnicí, která představuje Euklidovskou metriku - tedy prostou vzdálenost dvou neuronů.

Výstupní vrstva ale není omezena jen na dvoudimenzionální uspořádání, které je ale velmi vhodné pro vizualizaci výsledků. Pokud je to vhodné pro konkrétní řešený problém, může být výstupní vrstva uspořádána i lineárně (tedy bude vrstva jednorozměrná a definice sousednosti je vždy jen mezi dvěma prvky) nebo trojrozměrně (libovolná vzdálenost bodů v prostoru). Výstupní mapa může mít i sousedství definované tak, aby všechny body měly symetrické sousedství - body, které jsou na okraji vrstvy mohou mít definované sousedy na protější straně této vrstvy - v tu chvíli by tato vrstva tvořila pomyslný toroid. Topologie uvedená na obr. 5.9 obsahuje neurony, které mají asymetrické okolí.



Obr. 5.9: Architektura Kohonenovy sítě - SOM

Samotné okolí neuronu může být definováno jako ostré (funkce, která nabývá hodnot 1 resp. 0 pokud neuron náleží do okolí r daného neuronu, resp. nenáleží - obr. 5.12) nebo pomocí reálné funkce příslušnosti, která každému neuronu

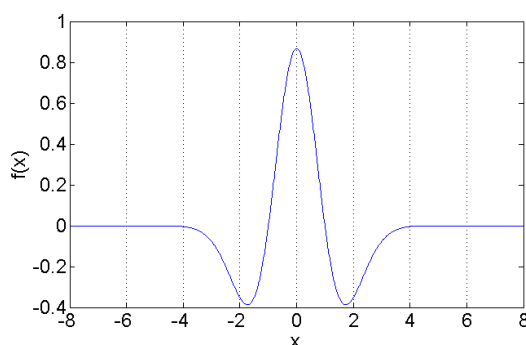


Obr. 5.10: Příklad možných definic uspořádání vzdáleností ve výstupní vrstvě SOM: a) čtvercové okolí, b) šestiúhelníkové okolí

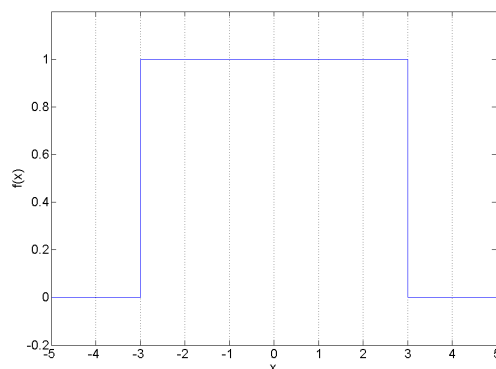
přiřazuje reálnou hodnotu, která určuje míru jeho příslušnosti k okolí daného neuronu. Touto funkcí bývá obvykle Gaussova funkce nebo funkce podobná wavelotové funkci mexického klobouku (Mexican Hat function - na obr. 5.11) definovaná jako:

$$\Theta(x) = \frac{2}{\sqrt{3\sigma\pi^{\frac{1}{4}}}} \left(1 - \frac{x^2}{\sigma^2}\right) e^{-\frac{x^2}{2\sigma^2}}, \quad (5.31)$$

tato funkce je průběhem podobná derivaci Gaussovy funkce.



Obr. 5.11: Průběh funkce mexického klobouku - rovnice (5.31)



Obr. 5.12: Průběh schodové funkce ostrého okolí pro $p = 3$

5.2.3.1 Učení Kohonenovy sítě

Učení SOM je učením bez učitele a je založeno na principu soutěžního učení (competitive learning). Síti se předkládají vstupní vektory (učící vzory) a mezi

neurony se na základě soutěže (definované soutěžní funkcí) vybere nejvhodnější neuron (neuron, který je nejbližší předkládanému vzoru). Vítězný neuron pak upravuje své váhy tak, aby jeho příští reakce na stejný podnět byla zesílena. Ostatním neuronům je také upravována váha, ale tato změna je určena funkcí příslušnosti k okolí vítězného neuronu. Obvykle tedy v nejbližším okolí vítězného neuronu také dochází ke zvýšení odezvy na předkládaný vzor, ale o menší hodnotu. Pro vzdálenější neurony může být pak citlivost na předkládaný vzor snižována. Tento přístup ke změnám vah neuronu odpovídá síti s funkcí mexického klobouku, která představuje funkci okolí (rovnice (5.31) a obr. 5.11). Tento přístup se nazývá laterální inhibice, který je inspirován neurobiologickými studiemi a je popsán například v [96]. Jednodušším přístupem je použití skokové funkce na obr. 5.12, kde jsou všechny neurony v okolí zcitlivěny na příslušný vzor a neurony mimo okolí nejsou změněny.

Průběh učení Kohonenovy sítě:

1. **Inicializace** - Inicializace vah w_{ij} na náhodné malé hodnoty, parametr učení ν a velikosti sousedství p (nebo jiné parametry funkce sousedství - σ pro Gaussovu funkci nebo funkci mexického klobouku).
2. **Předložení vzoru** - Na vstup sítě je předložen vzor, který má být učen: $\mathbf{x} = (x_1, x_2, \dots, x_n)$
3. **Vyhodnocení vzdáleností** - Pro všechny neurony j je vypočtena vzdálenost od učícího vzoru. Při použití Eukleidovské metriky:

$$d(j) = \sum_{i=1}^n (w_{ij} - x_i)^2. \quad (5.32)$$

4. **Výběr nejbližšího neuronu** - Je třeba najít neuron, který má minimální vzdálenost $d(j)$.
5. **Přízpůsobení vah neuronů** - Váhy neuronu j a všem z okolí Θ_j se upraví váhy podle:

$$w_{ij}(t+1) = w_{ij}(t) + \nu[x_i(t) - w_{ij}(t)]. \quad (5.33)$$

6. **Opakování** - Probíhá opakování učícího procesu (krok 2) pro další učící vzory, dokud neproběhne příslušný počet opakování nebo se nedosáhne požadované přesnosti.

Na začátku učení se učící parametr ν volí blízko jedné a postupně se snižuje směrem k nule. V průběhu učení se obvykle se snižováním učícího parametru zmenšuje i velikost okolí neuronů (zde záleží na použité funkci okolí a jejích parametrech).

Průběh učení u SOM lze velmi názorně vizualizovat. Každý neuron můžeme v ploše definovat pomocí váhových vektorů, které udávají jejich pozici. Pokud tedy zobrazíme všechny neurony v ploše včetně jejich propojení na sousedy, dostáváme

různě pokroucenou mřížku (po počáteční inicializaci na velmi malé ploše), která se v tomto prostoru po každé změně vah roztahuje. Cílem učení je dosáhnout co největšího roztažení mřížky bez různých překroucení.

5.2.3.2 Vybavování Kohonenovy sítě

Vzhledem k principu učení je zřejmé, že pro každý vstupní vektor v naučené síti vybavuje jen jeden neuron a to ten, který je podle definované funkce vzdálenosti nejbliž k předkládanému vzoru. Tato funkce vzdálenosti bývá nejčastěji definována jako Eukleidovská vzdálenost d mezi vektorem vah \mathbf{w} jednotlivých neuronů j a vstupním vektorem \mathbf{x} podle rovnice (5.32).

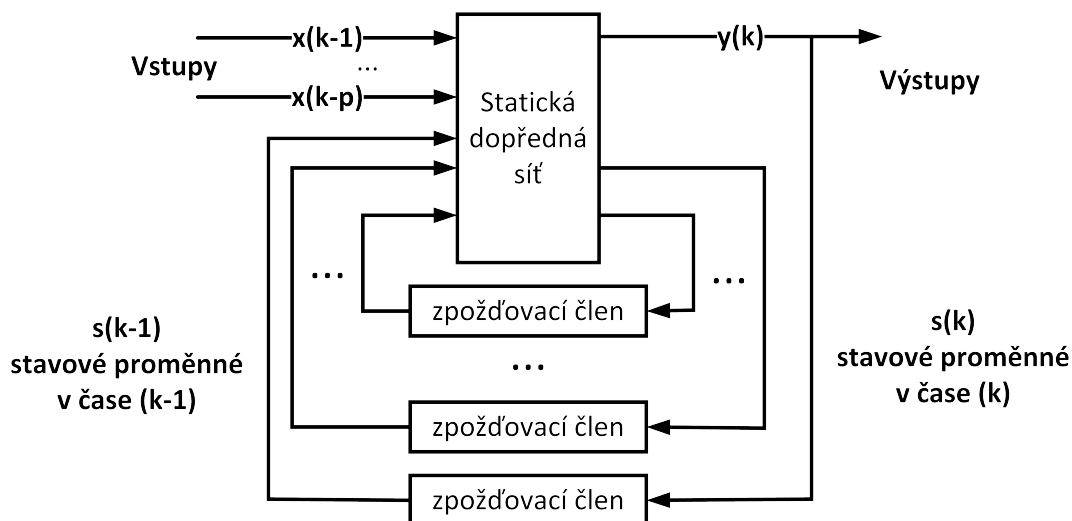
Je také možné, aby SOM pro některé vhodné problémy využívala pro vybavování i učení jinou metriku pro určení vzdálenosti. Tou může být i skalární součin vektorů \mathbf{w} a \mathbf{x} , který udává úhel mezi těmito vektory. Na základě tohoto principu vybavování je možné při vhodném způsobu vizualizace použít naučenou síť na detekci shluků ve vstupních datech. Je zřejmé, že podobné vektory budou aktivovat neurony v jedné oblasti. Tato vlastnost SOM sítí je velmi často využívána. Porovnání různých algoritmů shlukové analýzy bylo publikováno například v [97].

5.2.4 Rekurentní sítě

Rekurentní neuronové sítě jsou obdobou dopředných neuronových sítí resp. obsahují v sobě architekturu těchto sítí, ale navíc obsahují zpětné vazby. Tyto zpětné vazby zpravidla do sítě zavádějí časově závislý prvek. Lze tedy říci, že odezva neuronové sítě už nebude záviset jen na kombinaci vstupních hodnot, ale bude nějakým způsobem zohledněn i předchozí odezvu (nebo i jen část jejího výpočtu). V okamžiku, kdy síť obsahuje rekurentní spoje, je nutné tuto síť už učit celými sekvencemi vstupů a nikoli už jen jakkoli promíchanou množinou vstupních vektorů právě kvůli zavedenému časovému kontextu. Tyto sítě vyžadují jiný způsob učení, který ale stále může být založen na původním Back-propagation. Rozbor a algoritmus tohoto základního učení je uveden např. v publikaci [98].

Obecná rekurentní síť může být převedena do Nerrandovy kanonické formy (popsaná v práci [99]), která bude mít externí vstupy (ty mohou být časově závislé), časově závislý výstup, zpětnou vazbu stavových proměnných přes zpoždovací bloky a výkonná síť už bude běžná statická dopředná neuronová síť. Tato architektura je rozkreslena na obr. 5.13. V práci [100] byl tento přístup obhájen.

Určitém zjednodušením rekurentních sítí může být neuronová síť s časovými zpožděními (Time-delay Neural Network), kde jsou v síti zavedeny zpětné vazby z výstupu na vstup přes bloky zahrnující zpoždění - ty opět zajišťují síti časový kontext. Je to vlastně zjednodušení, kdy by v kanonické formě byly na nulu nastaveny váhy u neuronů, ke kterým jsou přivedeny stavové proměnné s z obr. 5.13. Obdobu těchto sítí bude využívána i později v této práci.



Obr. 5.13: Nerrandova kanonická forma

5.3 Aplikace neuronových sítí

Zde uvádím výčet některých aplikací výše uvedených neuronových sítí z oblasti elektromagnetického pole, meteorologie, ale i jiných oborů, které dobře demonstují využitelnost neuronových sítí. Zároveň je zde zřejmý posun k využívání neuronových sítí jako podpůrného prostředku pro jiné, již známé a dlouho využívané metody. Dopředné neuronové sítě

- Kombinace neuronové sítě a fuzzy-logiky pro vytvoření modelu predikce útlumu akustické vlny v atmosféře [101].
- Neuronový model predikce šíření radiového signálu v městské zástavbě [102].
- Zlepšení přesnosti meteorologického radaru pomocí dopředné neuronové sítě a měření množství srážek srážkoměrem v příslušné oblasti [101].
- Neuronová síť v kombinaci s genetickým algoritmem předpovídá množství vody v rozvodí řeky v severozápadním Íránu [103].

Kohonenovy samoorganizující mapy

- Shluková analýza na základě meteorologických dat pro další využití v meteorologii [104].
- Klasifikace druhu mraku na základě snímků z družice [105].
- Urychlení optimalizace patchové antény při použití sítě SOM [106].

Rekurentní neuronové sítě

- Krátkodobá předpověď spotřeby elektrické energie pro přenosové soustavy pomocí rekurentní neuronové sítě se spožděním v globálních zpětných vazbách [107].
- Několikadenní předpovídání rychlosti větru a budoucího výkonu pro několik farem větrných elektráren v Řecku pomocí rekurentních neuronových sítí [108].
- Předpovídání ceny akcií pomocí rekurentní neuronové sítě v kombinaci s různými přístupy k analýze hlavní komponenty (Principal Component Analysis) [109]

Kapitola 6

Odhad refraktivity pro pozemní radiové spoje

Tato kapitola je představuje nový způsob predikce chování dlouhých radiových spojů ve vnitrozemských podmínkách v závislosti na stavu počasí. Jako nástroj zajišťující predikci jsou zvoleny prostředky umělé/výpočetní inteligence. Jsou zde identifikovány fyzikální faktory ovlivňující šíření radiového signálu.

6.1 Dílčí cíle

Tato část práce má dva dílčí cíle:

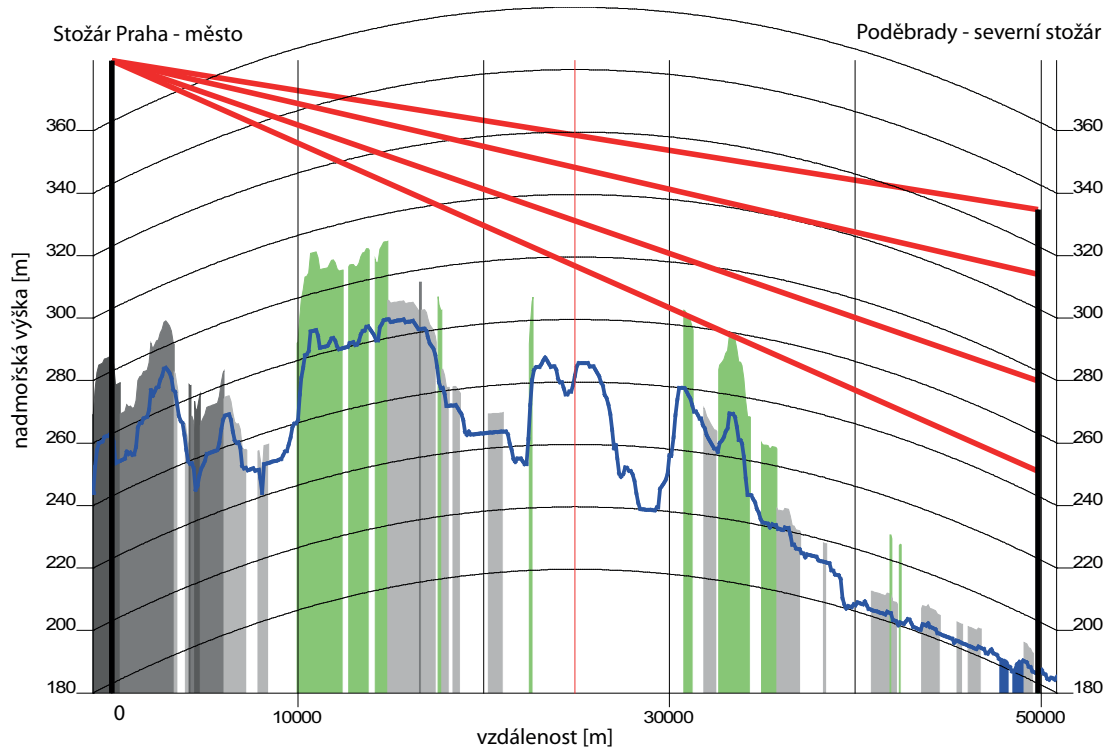
1. Ověřit možnost využití získaných dat k detekci stavu počasí na trase experimentálního spoje - remote sensing nejnižších troposférických vrstev a především jevů, které se v nich vyskytují
2. Vytvoření metodiky pro odhad gradientu refraktivity v troposférických vrstvách, kterými experimentální spoj prochází, a vyhodnocení kvality výstupu poskytovanou pro modely vytvořené užitím této metodiky.

Pro splnění obou cílů se předpokládá využití prostředků výpočetní inteligence, konkrétně umělých neuronových sítí.

6.2 Popis měřicího spoje

Pro tvorbu metodiky uvedé v této kapitole jsou použita data z experimentálního spoje [110], který není původně určen pro řešení zadaného úkolu. Lze říci, že vytvářená metodika vznikla zpětně na základě naměřených dat v předchozích letech. Nebyla tedy možnost jeho možnosti účinně měnit tak, aby lépe vyhovoval této práci. V době odevzdání této práce již tento experimentální spoj není v provozu. Experimentální měřicí linka je založena na kombinovaném měření meteorologických dat a útlumů experimentálních spojů, které jsou postaveny tak, aby spojnice

vysílače a přijímače protínala různé výškové hladiny. Nejnižší přijímač je situován tak, aby spojnice vysílače a přijímače procházela co nejnižše k terénní překážce (ve vzdálenosti cca 30 km od vysílače je to asi 25 m nad terénem pokrytým lesní vegetací) - ta zasahuje do Fresnelovy zóny spoje. Proto se předpokládá, že tato terénní vlna může mít velký vliv v případě superrefrakce, kdy může být nejnižší spoj zcela zastíněn touto překážkou. Celý terénní profil je zobrazen na obr. 6.1.



Obr. 6.1: Terénní profil experimentálního spoje s naznačenými jednotlivými spoji

Spoje jsou provozovány na frekvenci 10,671 GHz s amplitudovou modulací (modulační frekvence 1 kHz) a data s úrovněmi přijatého signálu jsou vzorkována po jedné sekundě. Vysílač je situován na budově Televizního vysílače Praha a stožár s přijímači je na bývalém vysílači Poděbrady - severním stožáru. Celková vzdálenost těchto dvou bodů je 49822 m. Nadmořská výška středu antény vysílače je 348,6 m a nadmořské výšky jednotlivých přijímačů jsou:

- 249,1 m
- 278,0 m
- 307,9 m
- 333,5 m

Nadmořská výška patky stožáru Poděbrady je 188,0 m nad mořem.

Meteorologické měření na straně vysílače neprobíhá přímo na vysílači, kvůli nedostatku místa pro umístění meteostanice, ale je situováno ve vzdálenosti cca 5,5 km od vysílače na pozemku Českého metrologického institutu. Na poděbradské straně spoje jsou umístěny tato čidla:

- Tlak - Vaisala BAROCAP PTB100A, přesnost v rozsahu -20–45°C je ± 1.5 hPa,

umístěn ve výšce 1,4 m

- Teplota a vlhkost - Vaisala HUMICAP HMP45D - přesnost $\pm 3\%$ rH, $\pm 0,2^\circ\text{C}$, umístěny na stožáru ve výškách od země:
 1. 5,11 m
 2. 27,61 m
 3. 50,26 m
 4. 75,86 m
 5. 98,26 m
 6. 123,86 m
- Srážkoměr
- Anemometr

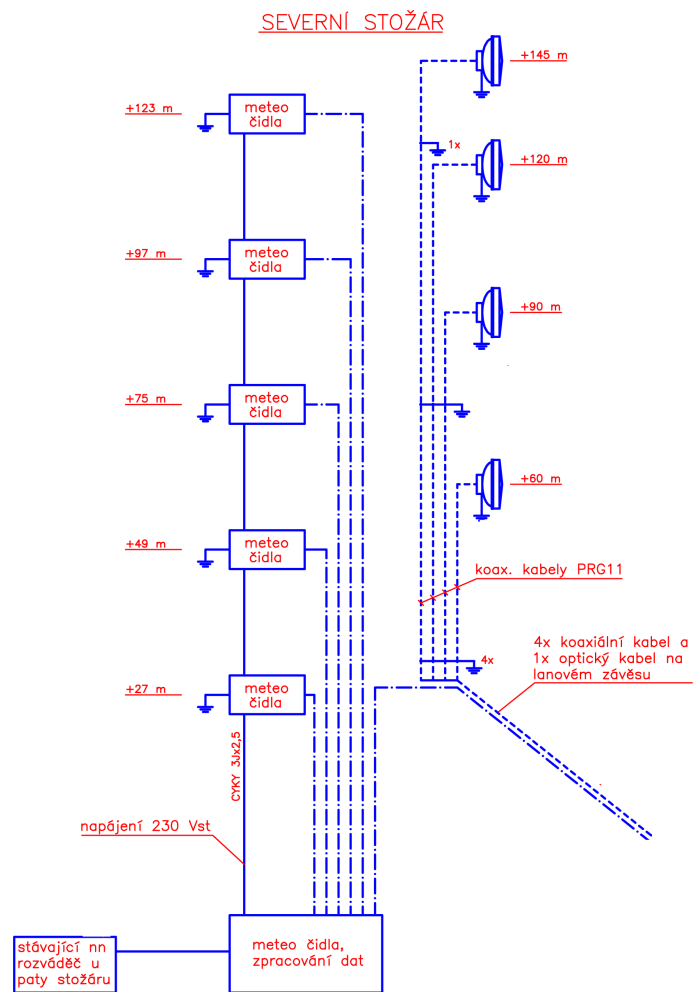
Celý stožár je schématicky rozkreslen na obr. 6.2 a fotografie tohoto stožáru jsou na obr. 6.3 a 6.4. Shromažďování dat probíhalo od roku 2008. Teplota, tlak a vlhkost jsou měřeny v jednodominutových krocích. Na straně vysílače je instalován senzor viditelnosti, jehož data jsou rovněž ukládána po jedné minutě. Množství srážek, směr větru a rychlost větru jsou měřeny pouze na straně přijímačů a jejich ukládání probíhá po 15 minutách. Abychom získali množinu dat, která bude dostatečně velká a bude v malých krocích, je potřeba data, která nejsou měřena v minutových intervalech, rozpočítat do těchto minutových intervalů.

6.3 Rozbor měřených dat

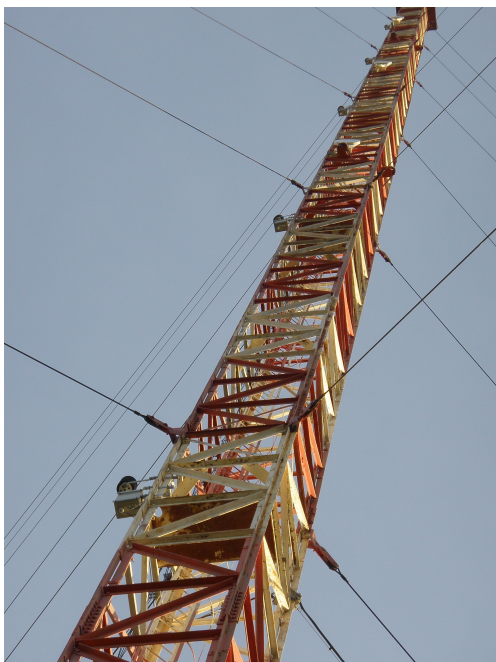
V prvním kroku proběhla ruční porovnání naměřených radiových a meteorologických hodnot. Bylo třeba zjistit, je-li vůbec v naměřených datech vidět vliv počasí. Bylo proto v meteorologických datech vybráno pouze období s měřitelným množstvím srážek. Pro tyto časové úseky byly v naměřených datech útlumů vyhledány příslušné hodnoty a ty jsou zobrazeny na obr. 6.5. Z průběhů je vidět, že útlum za deště roste až na hodnoty 5–20 dB v závislosti na síle deště, ale i na jiných meteorologických veličinách. Důležitou pozorovanou vlastností jsou fluktuace v hodnotách útlumů, které probíhají současně na všech patrech přijímačů.

Dále byly ručně analyzovány průběhy jednotlivých dnů (vždy 24 hodin, na grafech jsou na ose x minuty od začátku sledovaného období, které se může lišit na každém grafu), které měly definovaný průběh počasí. Na obr. 6.6 je průběh dne, kdy na začátku dne jsou krátké sněhové přeháňky. Atmosferický tlak pomalu roste v rozmezí 1008.8–1010.9 hPa. Vlhkost po ustání srážek klesá z 95% až na 80%. Během dne teplota kolísá v řádu 0.5°C za hodinu a v čase, kdy zapadne slunce, se teplota mění už jen velmi pomalu - tato závislost se projevuje malými fluktuacemi v počáteční a koncové fázi zobrazeného průběhu.

Na obr. 6.6 je zobrazen průběh s delším obdobím sněhových srážek během noci. Po ustání srážek přetrvávají stále stejné podmínky s malými změnami. Tep-



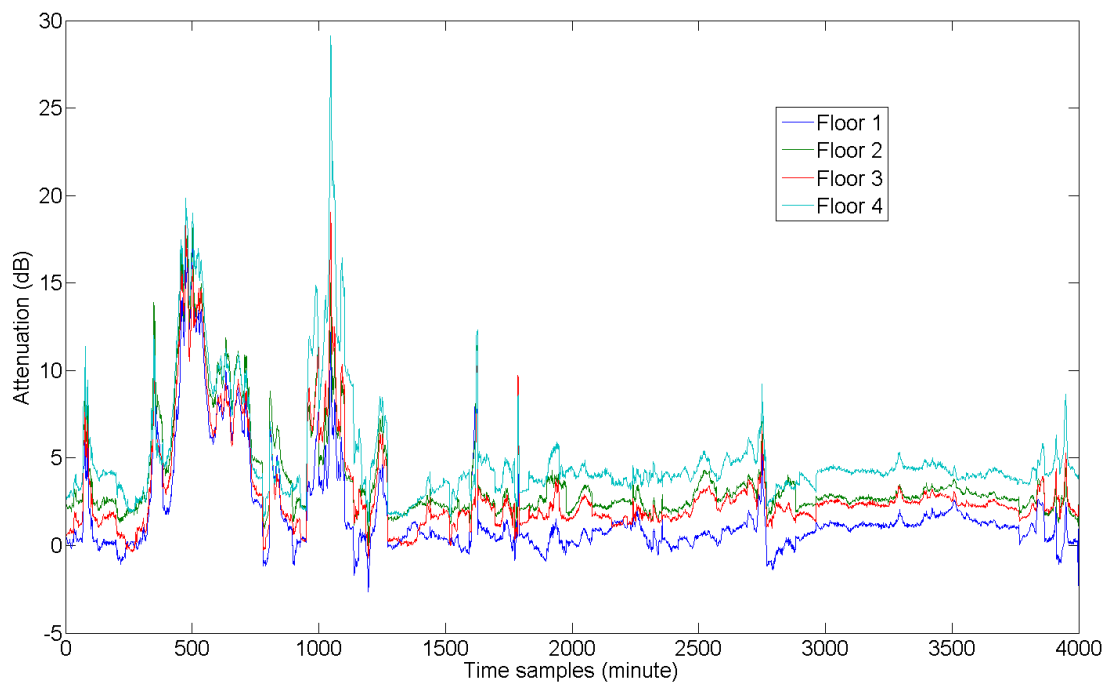
Obr. 6.2: Schéma uspořádání měřících zařízení na stožáru Poděbrady



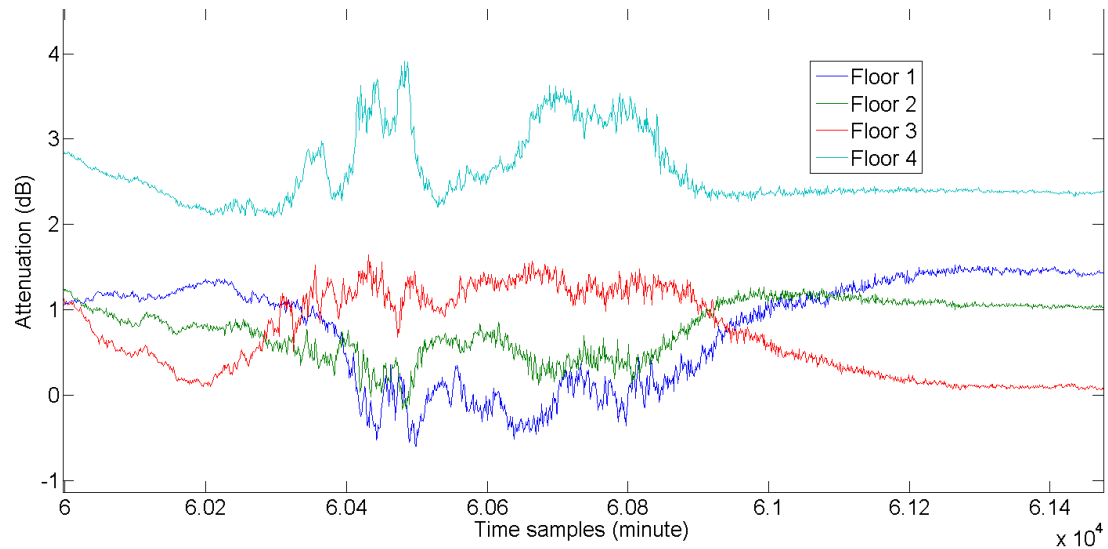
Obr. 6.3: Stožár Poděbrady s přijímači a meteosenzory - spodní pohled ze země



Obr. 6.4: Stožár Poděbrady - skříň se zařízením pro načítání dat a 1. čidlo teploty a vlhkosti

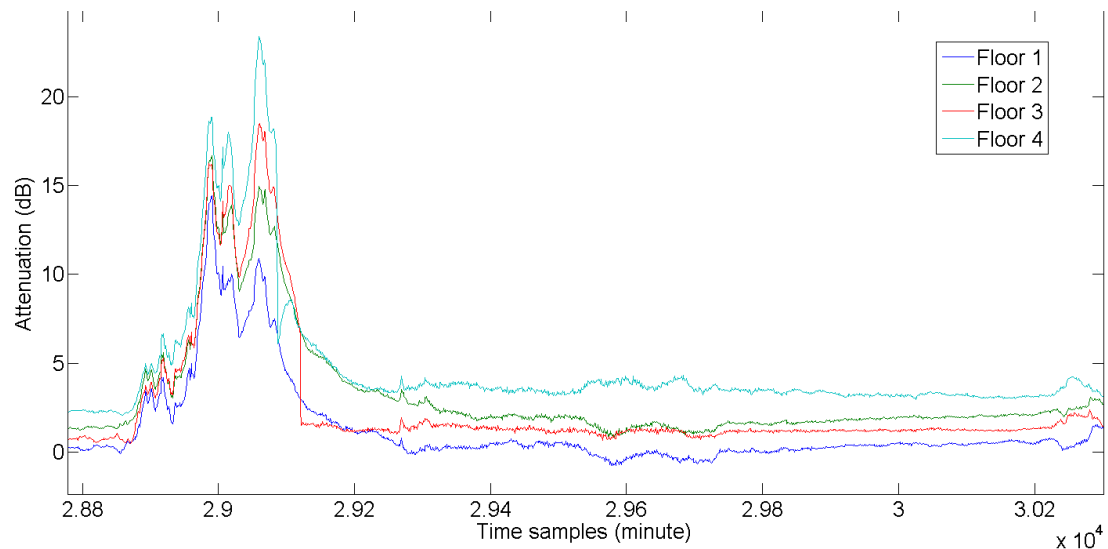


Obr. 6.5: Průběh útlumů měřicích spojů za deště



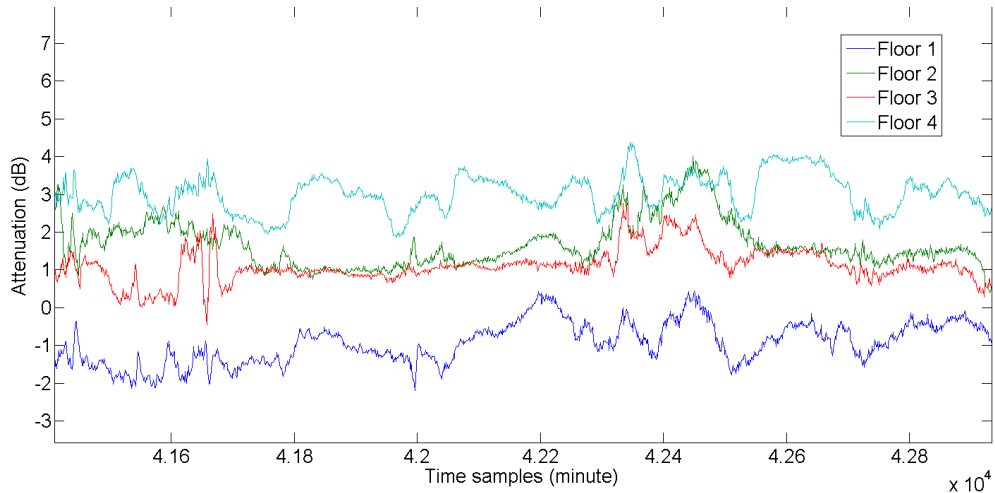
Obr. 6.6: Průběh radiových útlumů jednoho dne - 1/12/2007

lota postupně roste až do odpoledních hodin (rychlost cca $0,1^{\circ}\text{C}$ za dvě hodiny). Vlhkost pozvolna roste z 80% až na 90% v nočních hodinách. Tento průběh ukazuje, že při pomalých změnách počasí se i pomalu mění útlumy a jejich fluktuace jsou velmi malé. Naopak srážky mají velmi výrazný vliv na průběh útlumů.



Obr. 6.7: Průběh útlumů jednoho dne - 10/12/2007

Na obr. 6.8 je zobrazen den, kdy jsou na začátku a v polovině průběhu vidět změny způsobené dvěma lehkými přeháňkami. Během celého dne zůstávají stabilní teploty s malými a pomalými změnami, ale jsou pozorovány výrazné změny ve směru a rychlosti větru. Na průběhu je vidět, že se liší průběhy útlumů spojů v různých výškách - jeden průběh je na prvním a čtvrtém patře a naprosto odlišný průběh je na druhém a čtvrtém patře. Z toho lze odvozovat, že i změny ve větru by mělo být možné tímto experimentem podchytit.



Obr. 6.8: Průběh útlumů jednoho dne - 23/12/2007

6.4 Klasifikace množství srážek pomocí radiových spojů

Z analýzy naměřených dat v kapitole 6.3, je zřejmé, že srážky velmi silně ovlivňují bezdrátový spoj a vlastně způsobují jeho kompletní utlumení, ve kterém se podmínky šíření velmi rychle mění. Z toho důvodu jsem se pokusil vytvořit klasifikátor, který dokáže rozlišit množství srážek v atmosféře - nejlépe jen pomocí naměřených útlumů, případně při doplnění minimálního množství meteorologických dat. Tento klasifikátor může být později použit jako pomocný vstup pro budoucí odhad profilu refraktivity.

6.4.1 Vstupní data

Pro tento model byla shromážděna data popsaná v předchozích kapitolách za přibližně 6 měsíců. Nebylo možné použít větší množství dat, protože později došlo k výpadkům na měřicím zařízení a data byla nekonzistentní. Následně byly vytvořeny množiny dat v závislosti na množství srážek na srážkoměru a s pomocným vstupem z meteostanice Vaisala, který udává, jestli prší. Tyto množiny byly:

- CLEAR - množina dat, kdy nepršelo. To odpovídá cca 91,4% ze všech vzorků, kterých je přes 200000
- LIGHT - množina dat, kdy byl zaznamenán lehký déšť - 8,3% ze všech vzorků
- NORMAL - množina dat se středním množstvím deště - 0,2% ze všech vzorků
- HEAVY - velmi silný déšť. Tato množina obsahuje pouze 33 vzorků, to odpovídá 33 minutám silného deště v půlroce měření

Další množiny, které byly použity pro porovnání výsledků jsou:

- ALL - všechny vzorky za 6 měsíců měření s vynechanými nekonzistentními údaji (vlivem výpadku měření nebo jiných technických chyb)
- DAY - experimentálně vybrané množství 5000 vzorků tak, aby se v něm vyskytovaly vzorky ze všech skupin srážek v přibližně stejném zastoupení jako v množině ALL. Uvedené množství vzorků je vybráno v iterativním přístupu, kdy se zvyšoval počet prvků této a porovnávalo se rozložení těchto náhodně vybraných vzorků s množinou ALL. Při 5000 vzorcích bylo dosaženo vhodného množství vzorků tak, že v téměř každé náhodně vybrané skupině vzorků byla i data z množiny HEAVY. Tato množina svojí délkou odpovídá době přibližně 2,5 dne.
- MIX - 120 vzorků, z každé množiny 30 (kvůli malé velikosti množiny HEAVY nelze vybrat víc). Tato množina je vybrána kvůli možnosti porovnání vlivu konstrukce trénovací množiny na celkový výsledek.

6.4.2 Učící proces

Cílem je práce je vytvořit univerzální klasifikátor (na základě [7]). K tomu mají být využity neuronové sítě se dvěma nebo třemi skrytými vrstvami. Před provedením experimentu nemáme žádnou znalost o správném rozměru sítě, proto použijeme Baum-Hausslerovo pravidlo ([111]), které ale použijeme jen jako základní vodítko a vytvoříme několik sítí, které budou mít větší i menší rozměry než je tímto pravidlem navrhováno.

Všechny neuronové sítě mají ve vstupní vrstvě čtyři vstupní neurony a ve výstupní vrstvě je jeden výstupní neuron. Vstupní neurony slouží pouze jako rozvětřující prvky pro první skrytou vrstvu. Skryté vrstvy využívají neurony s aktivační funkcí ve tvaru unipolární sigmoidy. Stejnou aktivační funkci využívají i neurony výstupní vrstvy. Výstupní vrstva může být doplněna výstupní funkcí $f(x)$, která zajistí samotnou klasifikaci do 4 požadovaných množin:

$$f(x) = \begin{cases} \text{CLEAR} & \text{pro } x \in \langle 0; 0, 165 \rangle \\ \text{LIGHT} & \text{pro } x \in \langle 0, 165; 0, 5 \rangle \\ \text{NORMAL} & \text{pro } x \in \langle 0, 5; 0, \rangle \\ \text{HEAVY} & \text{pro } x \in \langle 0, 835; 1 \rangle \end{cases} . \quad (6.1)$$

Tato funkce bude ale použita až pro naučené sítě, aby nebyl ovlivňován algoritmus učení tímto zaokrouhlováním výsledku.

Samotné učení sítí (jejich rozměry jsou uvedeny v tab. 6.1) proběhlo v několika po sobě následujících cyklech za použití učícího algoritmu back-propagation. Učící parametry tohoto algoritmu se mezi těmito cykly snižovaly (podrobnosti už byly popsány v kapitole 5.2.2.1). Ukazatelem, který určuje, má-li se snížit učící

parametr algoritmu, je průběh MSE a jeho gradientu. V okamžiku, kdy změny v MSE dosahují po dlouhou dobu malých hodnot blížících se k nule, nastalo buď uváznutí v lokálním minimu nebo se učení dostalo do místa, kdy je učící parametr už příliš velký a síť osciluje. Uváznutí v lokálním minimu si v tu chvíli nemusíme všimnout, protože síť bude trénována ještě několikrát a bude vybrán nejlepší výsledek. Budeme tedy pouze snižovat učící parametr a tím se dostaneme blíže k minimu - učící parametr se ve většině případů snižoval až k hodnotě 0,0005. Pro největší síť bylo učení množiny DAY tak časově náročné, že nebylo možno dojít k použitelným výsledkům a to ani pokud bylo použito jiného učícího algoritmu jako je GDM back-propagation (gradient descent with momentum back-propagation).

6.4.3 Vyhodnocení učícího procesu

Nejlepší výsledky dosažené po proběhnutí všech učících cyklů jsou zobrazeny v tab. 6.1. Tyto výsledky byly dosaženy nejméně po čtyřech učení každé sítě s výjimkou dvou největších. Celková doba učení těchto sítí byla v řádu týdnů na běžném PC. V tabulce je vidět, že síť od velikosti 20x10 (20 v první a 10 v druhé skryté vrstvě) a všechny třívrstvé sítě jsou schopné se dobře naučit malou předkládanou učící množinou MIX. U trojvrstvé sítě 400x300x150 je už její učení tak časově náročné, že nebylo dále prováděno a jsou v tabulce nejlepší výsledky za simulace v řádu několika desítek hodin.

Na druhou stranu v případě překládní učící množiny DAY, jsou pro její učení nejvhodnější sítě o velikosti 20x10, 100x50, 60x30x15 a 100x50x25. Je vhodné podotknout, že vhodnější velikosti sítí jsou ty, které mají v poslední skryté vrstvě mezi 5 a 15 neuronů. Je tedy vhodnější pro tuto aplikaci používat menší množství neuronů v druhé skryté vrstvě než polovina neuronů ve vrstvě první.

Podrobněji jsou výsledky dosažené při trénování dvouvrstevných sítí zobrazeny na obr. 6.9 a 6.10 a pro trojvrstvé na obr. 6.11 a 6.12. Na těchto grafech lze podrobně vidět, jaké sítě jsou nejvhodnější pro použití při klasifikaci srážek. Tyto výsledky byly vytvořeny tak, že natrénované sítě byly předloženy příslušné množiny, které jsou uvedeny na ose x a byla vyhodnocena odezva. Při tomto vyhodnocení nebylo použito klasifikační funkce (6.1) připojené na výstup sítě. V případě, že se tato funkce použije, budou výsledky MSE lepší přibližně o 15% .

V grafech na obr. 6.9, 6.10, 6.11, 6.12 je vidět, že pro všechny sítě učené pomocí trénovací množiny DAY se objevují příliš velké chyby v odezvě na množinu HEAVY i NORMAL. To lze připisovat jejich malému zastoupení v této množině. Při porovnání výsledků sítí, které byly učené množinou MIX (která je navíc mnohem menší než DAY, ale obsahuje uniformě zastoupené jevy ze všech množin), lze z grafů vyčíst, že chyba v množině HEAVY se už nevyskytuje (to ale vzhledem k její velikosti bylo možné předem předpokládat a je tudíž toto nutno chápat spíše jako chybu při sestavování trénovací množiny) a chyby v ostatních množinách se vyrovnaly, ale pro skupiny NORMAL a LIGHT jsou na stejné úrovni.

Učící množina:		množina MIX		množina DAY	
Síť	Propojů	MIX	DAY	MIX	DAY
10x5	95	0.026	0.123	0.317	0.005
20x10	290	0.001	0.144	0.254	0.003
30x15	585	0.001	0.215	0.350	0.002
60x30	2070	0.000	0.179	0.362	0.001
100x50	5450	0.000	0.136	0.332	0.000
250x100	26100	0.000	0.122	0.332	0.000
500x250	127250	0.000	0.123	Nezdařilo se	
30x15x5	650	0.001	0.217	0.289	0.002
60x30x15	2505	0.000	0.174	0.302	0.001
100x50x25	6675	0.000	0.134	0.331	0.000
200x150x70	41370	0.000	0.161	0.331	0.000
400x300x150	166750	0.012	0.224	Nezdařilo se	

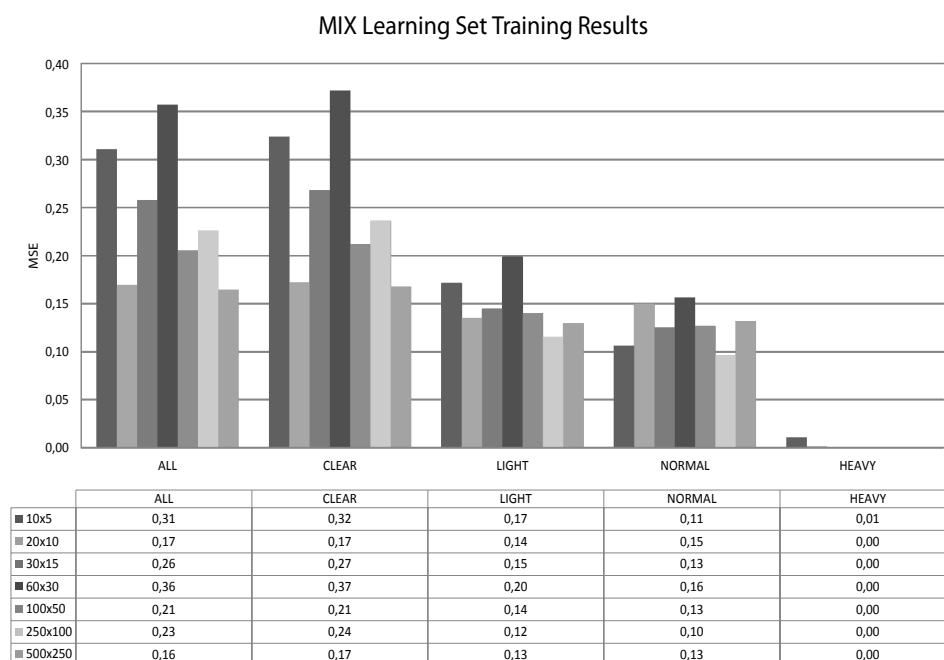
Tab. 6.1: Nejlepší MSE dosažené pro všechny sítě a učící množiny

Zhoršení tedy nastalo jen ve skupině CLEAR. Na základě těchto výsledků byly jako nejvhodnější sítě zvoleny 20x10, 100x50 a 100x50x25.

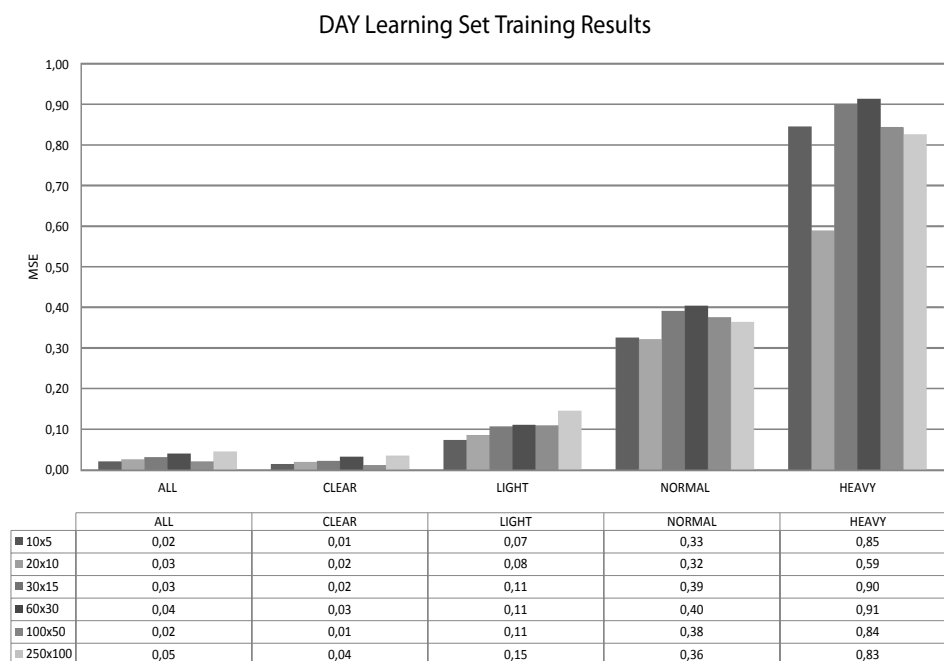
6.4.4 Dílčí závěry

Závěry plynoucí z této sekce jsou důležité pro další postup práce v oblasti odhadu refraktivity pomocí neuronové sítě. Nejsou tak důležité samotné velikosti neuronových sítí nalezené v této kapitole jako spíše tyto poznatky:

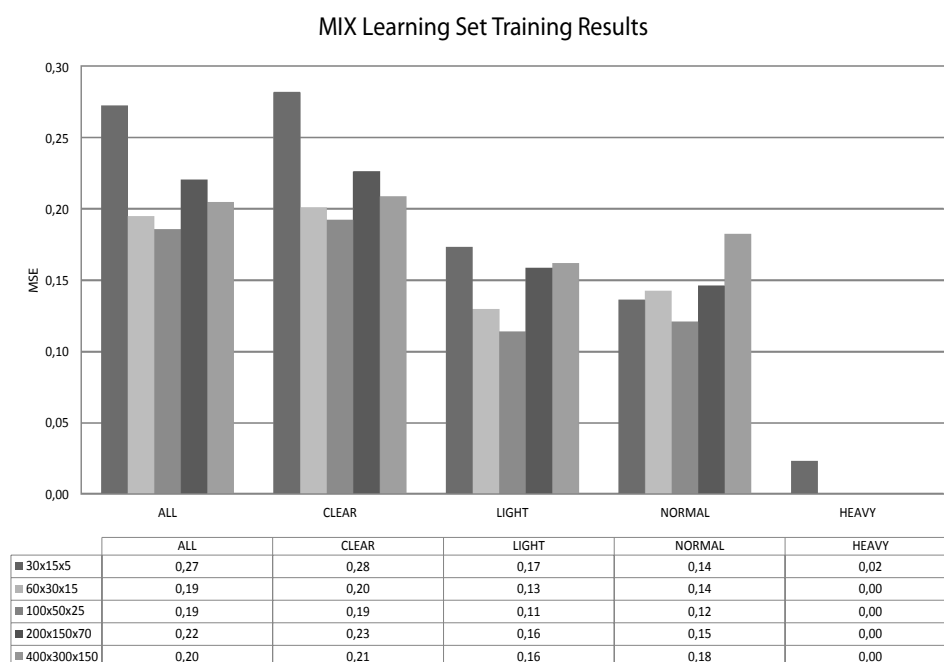
- Pro práci s meteorologickými daty a útlumy je třeba vhodným způsobem navrhovat trénovací množiny. Vzhledem k charakteru počasí je třeba vhodným způsobem zvýraznit jevy, které nastávají v krátkém čase z hlediska naměřených dat, ale ne příliš uniformě.
- Pro konstrukci neuronových sítí pracujících s meteorologickými daty jsou vhodnější sítě s větším množstvím neuronů ve skrytých vrstvách a druhá skrytá vrstva by měla mít menší počet neuronů než 1/2 první vrstvy.
- Je vhodné naměřená data doplnit o další meteorologická data z průběhu spoje, protože aktuálně měřená data nepokrývají celou délku spoje a nelze tak zohlednit například menší srážkové oblasti, které nejsou v dosahu pražské ani poděbradské meteorologické stanice. Tento požadavek je na použitím spoji těžko realizovatelný, ale pro případné další experimenty by k němu mělo být přihlédnuto.



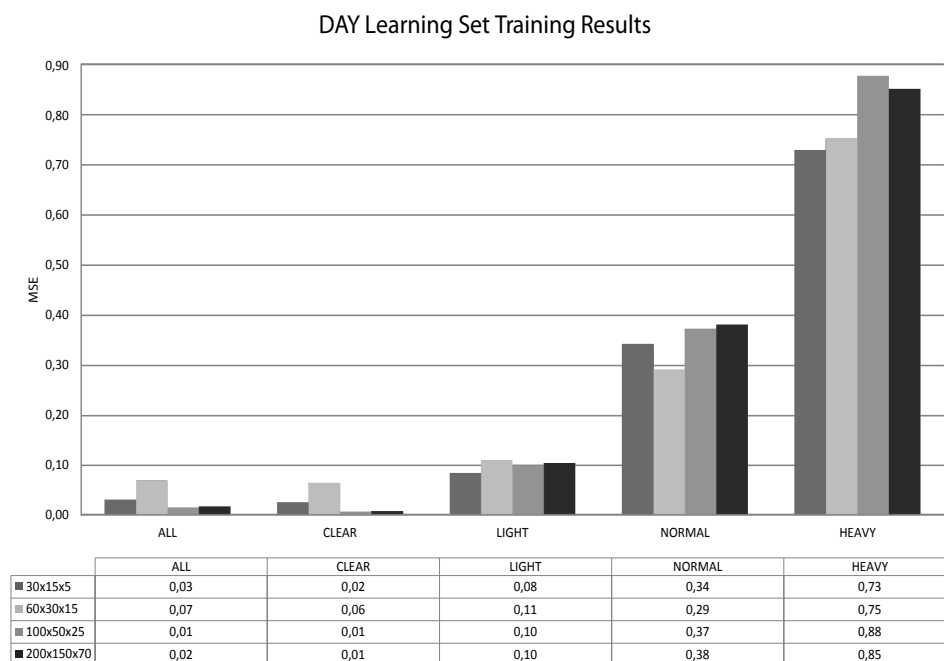
Obr. 6.9: Vyhodnocení MSE pro neuronové sítě se dvěma skrytými vrstvami a testovací množinu MIX



Obr. 6.10: Vyhodnocení MSE pro neuronové sítě se dvěma skrytými vrstvami a testovací množinu DAY



Obr. 6.11: Vyhodnocení MSE pro neuronové sítě se třemi skrytými vrstvami a testovací množinu MIX



Obr. 6.12: Vyhodnocení MSE pro neuronové sítě se třemi skrytými vrstvami a testovací množinu DAY

6.5 Konstrukce učicích množin pro meteorologické aplikace

Na základě závěrů popsaných v předchozím odstavci byly navrženy nové učicí množiny, které byly otestovány na stejně velkých sítích. Tyto množiny byly pojmenovány MIX, MIX2, MIX3 a jednotlivé srážkové třídy byly zastoupeny různě. Množina MIX zůstává stejná jako v předchozím odstavci, množiny MIX2 a MIX3 mají zvýrazněné zastoupení tříd CLEAR a LIGHT. Zastoupení jednotlivých tříd v učicích množinách je shrnuto v tab. 6.2. Všechny tyto množiny byly použity pro učení sítí 20x10, 30x15 a 100x50, první skrytá vrstva měla jako aktivační funkci bipolární sigmoidu (funkci $\tanh()$) a druhá vrstva už měla aktivační funkci ve tvaru unipolární sigmoidy (logistickou funkci). Toto učení probíhalo naprosto stejným způsobem jako v předchozím odstavci, bylo provedeno několikrát a byly vybrány nejlépe natrénované sítě.

Dešťová třída	Učicí množina		
	MIX	MIX2	MIX3
CLEAR	30	90	60
LIGHT	30	60	60
NORMAL	30	30	30
HEAVY	30	30	30

Tab. 6.2: Zastoupení jednotlivých tříd srážek v učicích množinách

Výsledky simulace sítí pro všechny naměřené hodnoty (třída ALL) jsou zobrazeny v tab. 6.3. Význam jednotlivých sloupců:

MSE chyba učení sítě

OK správně klasifikovaný podíl vstupních dat

ER1 podíl chybně klasifikovaných vstupních dat o jednu třídu - klasifikace dopadla do vedlejší třídy

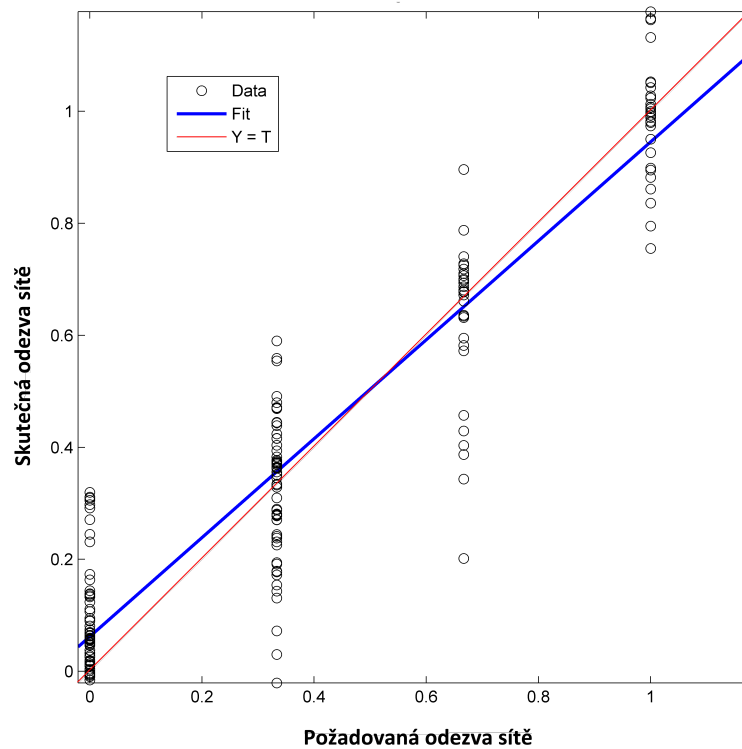
ER2 podíl chybně klasifikovaných vstupních dat o dvě třídy

ER3 podíl chybně klasifikovaných vstupních dat o tři třídy

Na základě zobrazených výsledků lze odvodit, že výše popsaný postup vytváření trénovací množiny zlepšil odezvu sítě o 15% až 20% ve správném odhadu množství srážek a třídy chyb ER2 a ER3 se zmenšily v některých případech o více než 10%. To je velmi výrazné zlepšení zvláště s přihlédnutím k faktu, že odhadované množství srážek není ostré a ve vstupních datech mohou být výrazné chyby způsobené nedostatkem informací o srážkách v průběhu spoje. Rozdíl požadovaných a získaných výsledků pro síť 100x50 a učicí třídu MIX2, kterou lze považovat za nejlépe natrénovanou (i přes to, že MSE není nejnižší) je zobrazena na obr. 6.13

Síť	Množina	MSE	OK	ER1	ER2	ER3
20x10	MIX	0.0098	35.3%	32.8%	17.3%	14.7%
	MIX2	0.0173	55.5%	32.5%	7.9%	4.1%
	MIX3	0.0204	48.2%	36.1%	11.9%	3.8%
30x15	MIX	0.0056	33.6%	29.2%	20.0%	17.3%
	MIX2	0.0257	50.3%	39.5%	6.9%	3.3%
	MIX3	0.0155	56.2%	26.5%	8.4%	9.0%
100x50	MIX	0.0307	31.0%	43.0%	13.6%	12.4%
	MIX2	0.0128	57.4%	26.3%	8.6%	7.7%
	MIX3	0.0108	47.7%	34.0%	11.6%	6.8%

Tab. 6.3: Kvalita klasifikace pro všechny učící množiny a neuronové sítě



Obr. 6.13: Příklad rozptylu dat pro nejlepší síť a trénovací množinu - 100x50 a množinu MIX2. Proloženo přímkou, která znázorňuje skutečný průběh (modrá) a červenou, která znázorňuje ideální závislost.

6.6 Odhad refraktivity a jejího výškového profilu

Jako první přiblížení se k odhadu výškového profilu refraktivity je použit jednodušší postup a tím je odhad gradientu refraktivity pro měřený spoj. Vzhledem k tomu, že nelze předem spoléhat na známé rozložení hodnot gradientu refraktivity v čase, byla pro tento postup vytvořena učící množina o velikosti 5000 vzorků

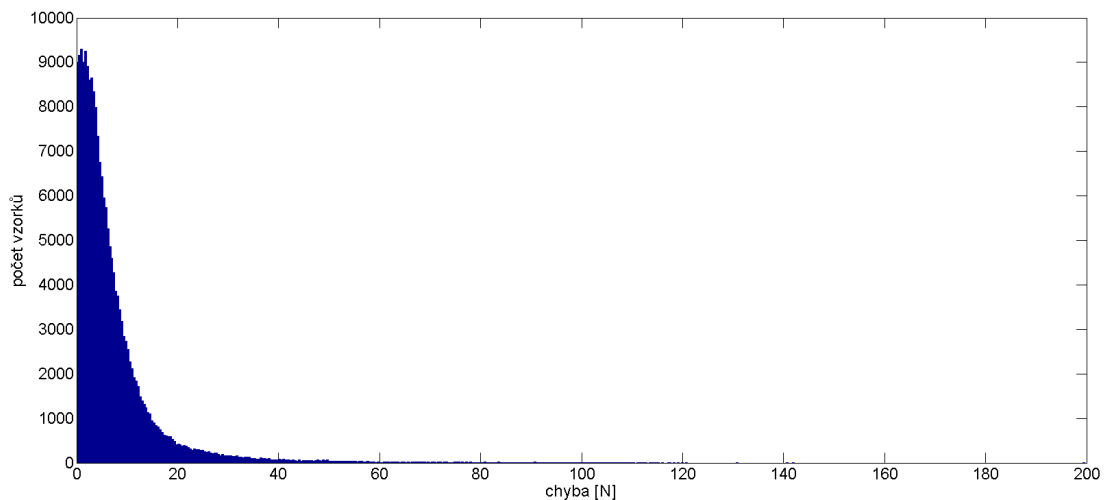
gradientu refraktivity. Gradient refraktivity byl vypočten na základě všech naměřených dat na stožáru s přijímači za použití vztahů (3.2), (3.4) a (3.5) včetně použití hodnot doporučených v ITU-R [13].

Pro ověření možností odhadu gradientu refraktivity byly využity už vytvořené sítě pro klasifikaci srážek jen s tím rozdílem, že jako aktivační funkce všech neuronů byla použita bipolární sigmoida, která lépe odpovídá řešenému problému. Gradient refraktivity se od průměrné hodnoty -40 N-units/km může pohybovat jak do kladných, tak i do záporných oblastí.

Výsledky trénování sítí o velikostech 20x10, 30x15 a 100x50 jsou zobrazeny v tab. 6.4 a histogram zobrazující chyby tohoto odhadu pro síť 30x15 je na obr. 6.14 .

Síť	MSE	Odchylka od naměřeného grad. refr. [N]				
		< 5	< 10	< 20	< 40	> 40
20x10	0.0456	53.0%	25.8%	14.7%	4.8%	1.7%
30x15	0.0436	53.7%	25.3%	14.6%	4.7%	1.7%
100x50	0.0472	52.4%	26.7%	14.3%	4.8%	1.8%

Tab. 6.4: Vyhodnocení přesnosti odhadu gradientu refraktivity



Obr. 6.14: Histogram chyby pro odhad gradientu refraktivity pro síť 30x15

6.6.1 Odhad profilu refraktivity

Vzhledem k množství meteorologických senzorů na stožáru v lokalitě Poděbrady (obr. 6.1 a 6.2) jsem se rozhodl pro výpočet průběhu gradientu refraktivity. V každém patře meteosenzorů je možné vypočítat refraktivitu a jejich rozdílem lze získat průběh gradientů refraktivity celkem v pěti vrstvách. Přejchod každé vrstvy odpovídá jednomu přijímači.

Pro řešení této problematiky byly zvoleny dopředné neuronové sítě se dvěma skrytými vrstvami. Vstupem jsou čtyři útlumy jednotlivých spojů a výstupem je pět gradientů refraktivity jednotlivých vrstev na stožáru Poděbrady. Rozměry sítí byly zvoleny 20x10, 30x18, 50x35 a 80x60. Větší rozměry sítí jsou už velmi těžko trénovatelné a podle předchozích výsledků není předpoklad, že by větší velikost sítě zlepšila přesnost odhadu. Jako aktivační funkce všech neuronů jsou zvoleny bipolární sigmoidy a pro učení bylo jako v předchozích případech zvolen algoritmus back-propagation, případně bylo v několika učících cyklech použito GDM back-propagation. Pro následující přehled výsledků byla vždy vybrána nejlepší síť.

6.6.2 Tvorba učících množin

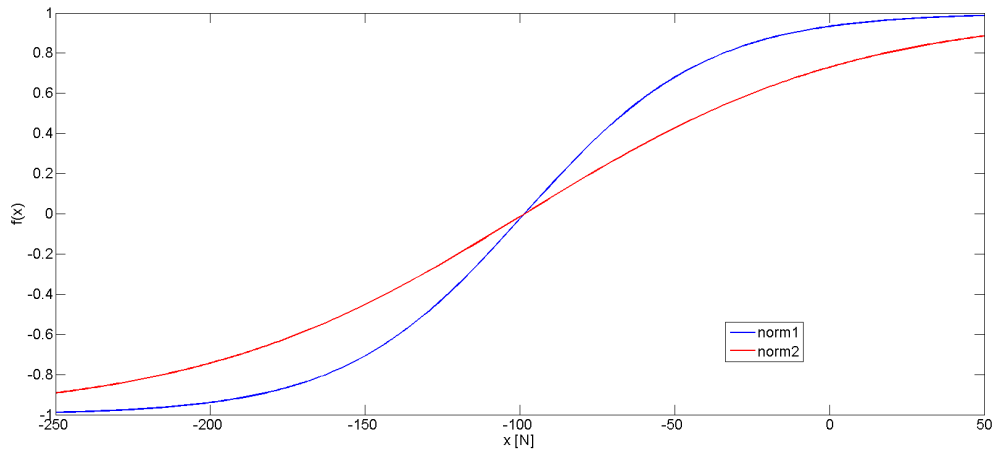
Důležitým vstupem do neuronové sítě je normalizace vstupních a výstupních dat. Jako vstupní data byly využívány útlumy. Tyto útlumy byly převedeny z logaritmické míry v dB na výkony a následně lineárně překódovány do intervalu $\langle -1; 1 \rangle$. Výstupní data odpovídající gradientům jsou rovněž převedena do intervalu $\langle -1; 1 \rangle$, ale tentokrát je k normalizaci použita vestavěná funkce Matlabu *tansig()* - stejná funkce jako aktivační funkce neuronů. Důvodem použití této funkce je její tvar. Pro normalizaci byly použity funkce, které jsou zobrazeny na obr. 6.15 :

$$\text{norm1}(x) = \text{tansig}\left(\frac{x + 98.5}{58.5}\right) \quad (6.2)$$

$$\text{norm1}(x) = \text{tansig}\left(\frac{x + 98.5}{106}\right) \quad (6.3)$$

Tyto funkce zajišťují, že veškeré vstupy refraktivity budou vždy ve správném intervalu. Dále jsou nastaveny tak, že v případě funkce (6.3) jsou hodnoty gradientu refraktivity -157 N/km rovny -0,5 a -40 N/km je rovno 0,5. Pro funkci (6.2) jsou hodnoty gradientu refraktivity -157 N/km rovny -0,75 a -40 N/km je rovno 0,75. To jsou hranice pro speciální troposférické efekty jako je vlnovodný kanál, subrefrakce a superrefrakce.

Učící množiny byly vytvořeny dvě (TS1 a TS2), každá za použité normalizační funkce s jiným průběhem ((6.2) a (6.3)). Vzhledem k tomu, že průměrná hodnota gradientu refraktivity pro střední Evropu je -40 N/km, je vybráno 350 vzorků z intervalu $\langle -10; 60 \rangle$ a 150 vzorků ze zbylého intervalu. Za použití těchto vzorků byly učeny všechny sítě a vyhodnocena jejich výkonnost.



Obr. 6.15: Průběh funkcí (6.3) a (6.3) použitých k normalizaci

6.7 Vyhodnocení kvality odhadu profilu refraktivity

Výsledky odhadu gradientu refraktivity pro jednotlivé učící množiny a celkový gradient (gradient od prvního meteočidla na zemi až po nejvyšší čidlo na stožáru) je uveden v tab. 6.5. Odezva **True** znamená, že gradient refraktivity se od požadované hodnoty neměnil o více než 10 N/km (to je hodnota menší než předpokládaná nejistota měření, která byla publikována v [112]), naproti tomu **False** znamená rozdíl o více než 100 N/km (přibližně rozdíl mezi vlnovodným kanálem a subrefraktivitou). Z uvedených údajů vyplývá, že využití obou normalizačních funkcí je pro tuto aplikaci vhodné a přesnost odhadu se pro všechny použité sítě pohybuje kolem 75%.

Množina	20x10		30x18		50x35		80x60	
	True	False	True	False	True	False	True	False
TS1	78,70%	1,10%	74,60%	1,00%	75,50%	0,70%	74,70%	0,50%
TS2	77,90%	1,50%	74,10%	0,60%	72,40%	0,70%	72,90%	0,70%

Tab. 6.5: Vyhodnocení odhad celkového gradientu refraktivity pro všechny trénovací množiny

Tab. 6.6 uvádí odhad profilu refraktivity pro každou jednotlivou vrstvu při použití normalizační funkce (6.3). Je zde vidět, že odhad refraktivity v jednotlivých patrech spoje se pohybuje od 70% do 95% a pro tuto výkonnost je dostačující neuronová síť o rozměru 20x10. Při porovnání s výsledky v tab. 6.6 lze vyvodit závěr, že normalizační funkce (6.3), která má větší sklon lineární části průběhu, poskytuje mírně přesnější výsledky pro síť 30x18. Oba tyto závěry naznačují, že metoda je použitelná pro odhad refraktivity na základě útlumů.

Dále je třeba vyhodnotit, s jakou přesností je odhadnut výškový profil refraktivity pro všechny měřené vrstvy. Rozložený výškový profil se skládá z pěti gra-

Vrstva	20x10		30x18		50x35		80x60	
1	60,20%	2,30%	60,00%	2,70%	54,00%	4,10%	39,10%	10,80%
2	84,10%	0,90%	83,50%	2,00%	79,20%	2,00%	77,40%	2,60%
3	70,00%	1,70%	67,60%	2,00%	64,00%	3,70%	53,00%	10,30%
4	86,20%	2,30%	81,70%	2,10%	72,60%	8,10%	65,50%	6,70%
5	90,60%	1,90%	88,40%	1,00%	84,40%	2,40%	70,70%	6,20%

Tab. 6.6: Vyhodnocení odhad jednotlivých pater gradientu refraktivity pro učicí množinu TS1

Vrstva	20x10		30x18		50x35		80x60	
1	67,70%	0,80%	67,90%	0,20%	53,40	5,40%	62,70%	1,80%
2	85,20%	0,60%	90,40%	0,00%	79,00	3,20%	85,80%	0,50%
3	71,00%	1,00%	73,00%	0,00%	64,70	2,90%	67,00%	1,90%
4	88,60%	0,60%	91,30%	0,00%	82,60	2,30%	83,80%	0,80%
5	91,40%	0,30%	95,00%	0,00%	84,20	2,70%	88,00%	1,10%

Tab. 6.7: Vyhodnocení odhad jednotlivých pater gradientu refraktivity pro učicí množinu TS2

dientních vrstev. Budeme předpokládat, že správný gradient refraktivity (**True**) je ten, který se od změřeného neliší v žádném patře o více než 10 N/km. Naproti tomu chybný gradient (**False**) je ten, který se liší v o více než 10 N/km ve čtyřech nebo pěti vrstvách. Jak lze vyčíst z tab. 6.8, nejlepší přesnost odhadu refraktivity byla pro síť 30x18 a trénovací množinu s normalizační funkcí (6.3).

Množina	20x10		30x18		50x35		80x60	
TS1	34,80%	4,50%	33,70%	6,10%	26,40%	9,80%	20,10%	20,10%
TS2	39,60%	3,30%	41,50%	0,70%	30,60%	8,20%	34,90%	4,40%

Tab. 6.8: Vyhodnocení celkového odhadu gradientu refraktivity pro obě trénovací množiny

6.8 Shrnutí použité metodiky

Samotná metodika pro tvorbu neuronového modelu odhadu refraktivity atmosféry se skládá z několika kroků, které byly podrobněji popsány v předchozích odstavcích. Součástí metodiky jsou doporučení, která je vhodné dodržet, aby bylo dosaženo co největší přesnosti. Metodika se skládá z následujících kroků:

1. Příprava a konverze měřených dat – odstavec 6.6.2.

2. Návrh neuronové sítě – probíhá experimentálně v několika krocích a musí být zpětně vyhodnocen podle odstavce 6.7.
3. Tvorba učících množin – odstavec 6.6.2 a 6.5.
4. Učení neuronové sítě – odstavec 6.4.2.
5. Vyhodnocení výstupů neuronové sítě – odstavec 6.7.

Pro zpracování naměřených dat lze shrnout následující doporučení, která je vhodné dodržet při použití metodiky, která byla popsána v přechozím textu:

1. **Normalizace vstupních dat.** Tento druh dat má tu vlastnost, že největší množství dat se nachází v okolí střední hodnoty, ale tato data ovlivňují samotné zpracování neuronovou sítí jen ve velmi malé míře. Je tedy vhodné použít už normalizační funkci, která tuto vlastnost zohlední při převodu vstupních hodnot do příslušného intervalu nutného pro zpracování pomocí neuronové sítě.
2. **Konstrukce učících množin.** Ze stejného důvodu jako v předchozím bodě je nutno učící množiny vytvářet tak, aby se zvýraznil výskyt hodnot mimo 90% kvantil. Není ale vhodné vytvořit zcela uniformní zastoupení, je-li klasifikováno jen několik tříd.
3. **Velikost neuronové sítě.** Velikosti použité sítě se pohybovaly ve velikostech 20x10 až 30x18 a větších. Nebyl pozorován zásadní vliv větších sítí na kvalitu odhadu refraktivity ani množství srážek. Naopak se při trénování objevovaly problémy s časovou náročností učení velkých sítí a pak s přeučení sítě.

6.9 Dílčí závěr

V kapitole byla představena metodika, jak přistupovat k naměřeným meteorologickým datům, aby je bylo možné využít pro zpracování neuronovými sítěmi. Zároveň byl naznačen přístup k diagnostice množství srážek pomocí experimentálního mikrovlnného spoje. Bylo dosaženo přesnosti klasifikace čtyř úrovní srážek s přesností 60%.

Byly vyhodnoceny výstupy poskytované metodikou. Přesnost těchto výstupů je postačující pro orientační odhady gradientu refraktivity, kdy přesnost dosahuje 70-80%. Přesnost odhadu výškového průběhu gradientu refraktivity klesá k 50%, to je ale dáno kombinací nepřesnostmi měření i nepřesností výstupů neuronové sítě a zároveň i vlivy, které nebylo možno měřit pomocí experimentálního spoje.

Přesnost dané metod je ovlivněna především dostupností vstupních dat a to jak časovou tak i prostorovou. Jak bylo zjištěno už na počátku, například déšť

způsobuje velké ztlumení spojů a během toho vznikají velké fluktuace v útlumech, které jsou jen těžko odhadnutelné. Dalším problémem je, že jediné měření profilu refraktivity probíhá jen na konci linky. Vzhledem k dynamice změn počasí na tak dlouhém spoji tak získáváme jen jakési zprůměrování profilu refraktivity, resp. známe jen výsledný efekt průběhu refraktivity.

Kapitola 7

Návrh metodiky predikce vlivu počasí na FSO spoje

7.1 Dílčí cíle

Cílem této části práce je návrh metodiky k vytvoření komplexního modelu vlivu počasí na kvalitu optického spoje (na útlum prostředí), tedy nejen v analýze jednotlivých atmosferických vlivů zvláště. K tomu mají sloužit lehce dostupná meteorologická čidla, která jsou umístěna v blízkosti sledovaných spojů. Sledovanými meteorologickými údaji jsou teplota, tlak, množství srážek, směr a rychlost větru. Není cílem jednotlivé vlivy analyzovat, výsledný model by měl fungovat jako samostatný modul („black box“), který se navrženým postupem naučí na konkrétní data a dále bude schopen na základě meteorologických dat predikovat chování FSO spoje. Jako vhodný prostředek pro tento úkol byly zvoleny neuronové sítě - předpoklad jejich vhodnosti byl učiněn na základě závěrů z jejich využití při návrhu systému pro odhad refraktivity v předchozí kapitole.

Vzhledem k tomu, že celý postup vytváření modelu je prováděn pro dva bezdrátové spoje, kdy každý poskytuje údaje o velikosti útlumu v jiném kvantovacím kroku, bude vyhodnoceno, jak jemné informace o útlumu navrhovaná metoda vyžaduje.

Zároveň tato práce slouží jako ověřovací studie metodiky popsané v kapitole věnované radiovým spojům, kde byly navrženy důležité postupy pro práci s naměřenými atmosferickými daty. Právě tyto postupy budou využity ke konstrukci vhodné architektury neuronové sítě, která dokáže predikovat chování. Původní metodika byla zaměřena na odhad profilu refraktivity pomocí bezdrátových spojů, který na tomto experimentu s FSO není aplikovatelný.

7.2 Popis experimentu

Pro měření byly použity dva FSO spoje, které jsou umístěny na budově FEL, ČVUT v Praze a na okolních budovách kampusu. První spoj, ze kterého byla

získávána data po vývoj modelu, má délku přibližně 120 m a je umístěn na ochozech v osmém patře na dvou blocích budovy Technická 2 - tedy přibližně 30 metrů nad zemí. Budova je zděná s velkým množstvím otevíratelných oken (to může způsobovat určité teplotní turbulence v okolí hlavic), na zemi je parkoviště a několik vzrostlých stromů, které nijak nezasahují do spoje. Tento spoj využívá starší 100 Mbps technologii WaveBridge 500 s vlnovou délkou 800-900 nm, šířkou paprsku 17 mrad. Spoj je orientován v severozápadním směru a z těchto směrů není zastíněn jinými budovami. U tohoto spoje byly pozorovány výkyvy v důsledku osvětlení hlavic spoje vycházejícím sluncem.

Druhý spoj je přibližně 450 m dlouhý a je orientován v západním směru. Pro tento spoj je využita novější čtyřsvazková technologie LightPointe FlightStrata-G s rychlostí 1000 Mbps. Zařízení pracuje na vlnové délce 850 nm se šířkou paprsku 2 mrad a výstupním výkonem 24 mW. Citlivost přijímače je -45 dBm, dynamický rozsah 34 dB. Výrobce doporučovaná maximální délka spoje je 1000 m stejně jako u prvního spoje.

Spoje jsou doplněny dvěma meteorologickými stanicemi Anemo WS981, umístěnými na střeše budovy v místě, kde jsou instalovány i hlavice FSO spojů. Jedna stanice shromažďuje data o teplotě, vlhkosti, atmosferickém tlaku, množství srážek (k tomu je použit vyhřívaný člunkový srážkoměr s rozlišením 0.1 mm a sběrnou plochou 500 cm²) a rychlosti a směru větru (anemometr AN 955C). Druhá stanice měří teplotu, atmosferický tlak a množství srážek starším srážkoměrem. Tato stanice je nově vybavena senzorem dohlednosti Vaisala PWD20 pro měření viditelnosti do 20 km - toto čidlo ale nespĺňuje stanovené podmínky pro model, kvůli tomu, že to není jednoduše dostupné a levné čidlo. Proto nebude využito v modelu. Všechna meteorologická data jsou měřena po jedné sekundě s výjimkou pomalu se měnících veličin jako je množství srážek, kdy je měření po sekundě prakticky neproveditelné. Pro tento případ byly provedeny přepočty z patnáctiminutových vzorků.

Data o přijímané úrovni signálu z jednotlivých hlavic FSO jsou rovněž ukládána po jedné sekundě. Hlavice WaveBridge poskytuje signál RSSI, který je pomocí naměřené charakteristiky převeden na dB. Hlavice FlightStrata poskytuje úroveň přijímaného signálu pouze v 10 úrovních, tyto úrovně jsou také přepočteny na dB. Důvodem je to, aby modely pro oba spoje měly stejný typ vstupů a výstupů.

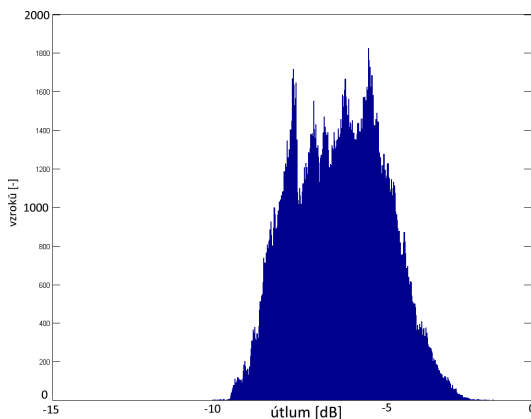
7.3 Analýza vstupních dat

První analýza vstupních dat pomocí SOM sítí byla publikována v [113]. Motivací k analýze vstupních dat byl pokus o zmenšení velikosti množiny vstupních dat tím, že v nich budou nalezeny shluky dat, případně silně korelované vstupní veličiny. Korelace vstupních veličin byla provedena statisticky pomocí výpočtu matice korelací všech vstupních veličin. Silně korelovanými veličinami byly pouze

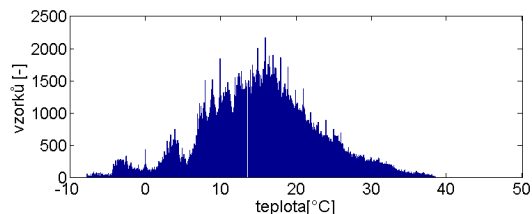
teplota, tlak a vlhkost z obou meteostanic, kdy koeficient korelace dosahoval hodnoty 1,00. Z toho důvodu nebyla tato (vlastně duplicitní) data z jedné z meteostanic dále využívána. Mezi ostatními vstupy nebyla nalezena žádná statisticky významná korelace.

7.3.1 Shluková analýza pomocí SOM sítě

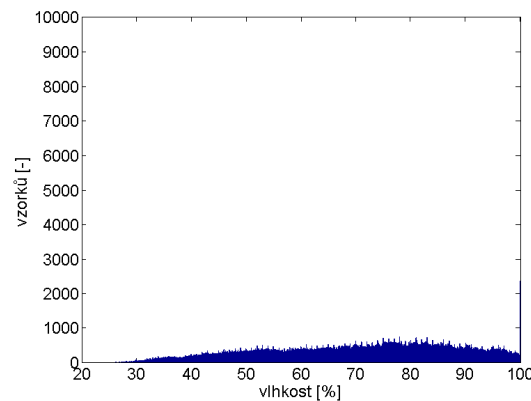
Z naměřených dat byly náhodně vybrány vzorky meteorologických dat (teplota, vlhkost, množství srážek) a útlumu s cílem ověřit výskyt shluků, který není zřejmý a nebyl nalezen pomocí statistiky - korelace. Množin vzorků bylo celkem 5 pro velikosti od 1000 do 5000 vzorků. Na ověření byly použity SOM sítě s hexagonálním uspořádáním a rozměrech od 10x10 do 30x30. Pro lepší přehled o rozložení naměřených dat jsou na obr. 7.1, 7.7 a 7.3 zobrazeny histogramy pro teplotu, vlhkost a útlum a na obr. 7.4 je zobrazen průběh srážkové činnosti v čase. Tyto histogramy svým průběhem odpovídají i později používaným datům, která byla získána za delší časové období.



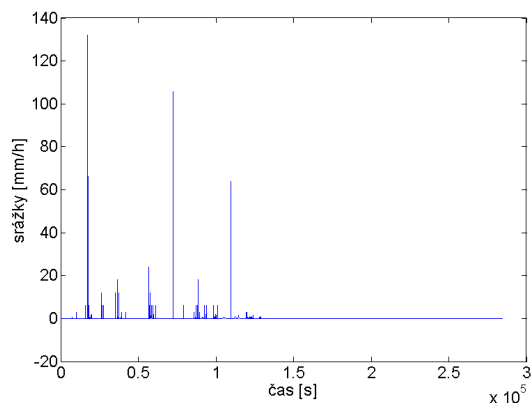
Obr. 7.1: Histogram naměřeného útlumu



Obr. 7.2: Histogram naměřené teploty

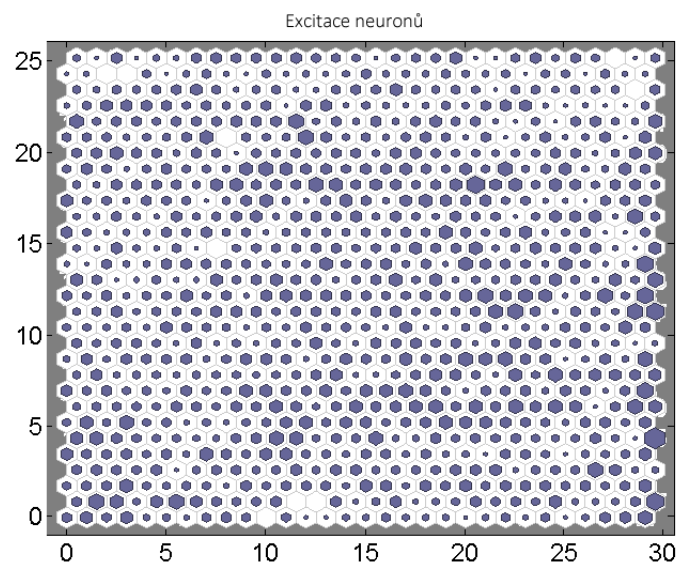


Obr. 7.3: Histogram naměřené vlhkosti

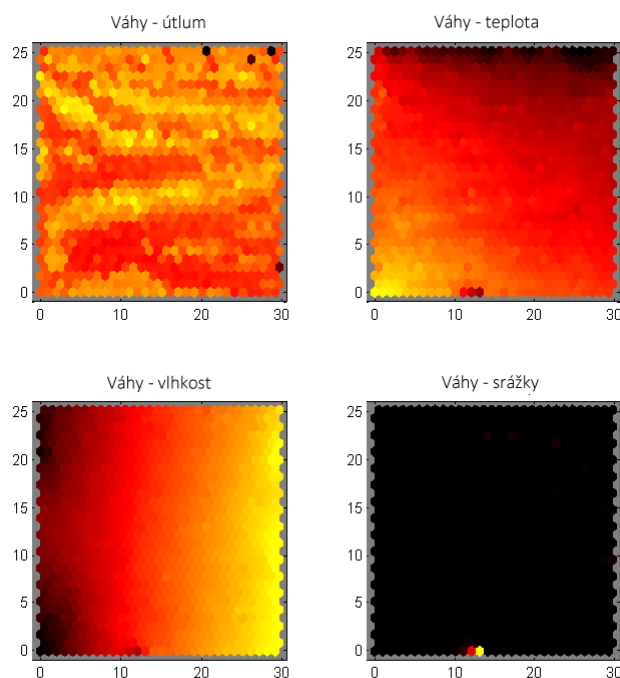


Obr. 7.4: Časový průběhu srážek

V žádném z provedených učících cyklů nebyl nalezen dostatečně silný shluk. Pro ilustraci uvedu výsledky pro učení sítě 30x30 a učící množinu o velikosti 2000 vzorků. Na obr. 7.5 jsou zobrazeny míry excitace jednotlivých neuronů - čím větší je modrá výplň hexagonálního neuronu, tím víc vstupů tento neuron excituje. Podle rozložení excitací lze usuzovat, že shluky dat nejsou významné. Doplňkovou vizualizací k této síti jsou zobrazeny mapy výstupních vah na obr. 7.6. Je zde zobrazeno, které neurony jsou daným vstupem silněji ovlivňovány (tmavá barva značí vyšší váhu). V případě, že by některá z map byla podobná jiné, lze mezi těmito vstupy předpokládat nějakou vazbu. To se v tomto případě rovněž nepotvrdilo.



Obr. 7.5: Zobrazení míry excitace jednotlivých neuronů



Obr. 7.6: Mapy vah pro všechny vstupy

7.4 Návrh metodiky modelování atmosferických vlivů na FSO spoj

Metodika, která je použita pro návrh modelu chování FSO spoje, se skládá z řady následných kroků, které budou v dalším textu podrobněji popsány. Některé části metodiky již byly popsány v kapitole věnované odhadu troposférické refraktivity, protože povaha vstupních dat byla obdobná. Pro vytvoření vhodného modelu FSO je třeba provést následující kroky:

1. **Naměření meteorologických dat a dat přijatých úrovní signálu (případně přepočtených na útlum).** Tato data by měla být měřena v co nejmenších časových intervalech. Toto tvrzení je podpořeno naměřenými hodnotami rychlosti scintilace v [73], [52] a především v [74].
2. **Normalizace naměřených dat.** V předchozích odstavcích bylo zobrazeno typické rozložení naměřených meteorologických dat. Tyto hodnoty je vždy potřeba převést do intervalu vhodného pro zpracování neuronovou sítí. Tento krok je důležitý pro získání dobré výsledné přesnosti modelu.
3. **Výběr vhodné architektury neuronové sítě.** Tento krok je obecně náročný, protože podmínky experimentů se liší a vzhledem k povaze neuronové sítě není možné dopředu určit přesný rozměr neuronové sítě, ale lze

poskytnout doporučení podložená srovnáním různých velikostí a architektur. Předpokládá se, že při zavedení časového kontextu do neuronové sítě se kvalita modelu zlepší.

4. **Příprava učících množin.** Krok který už byl podrobně analyzován v předchozí kapitole. Stejný navržený postup bude použit i v modelu FSO. Tento krok je prováděn až po vytvoření neuronových sítí, protože tyto sítě se mohou lišit počtem vstupů.
5. **Učení neuronové sítě.** Učení sítě v této metodice je prováděno algoritmem back-propagation, případně jeho modifikacemi. Byl zvolen stabilní algoritmus, který ale nemusí být vždy časově optimální. Postup učení už byl rozebrán v kapitole věnované odhadu refraktivity a bude dodržen.
6. **Vyhodnocení kvality modelu.** Na základě všech naměřených dat bude proveden odhad a budou porovnány odezvy všech navržených neuronových sítí. Výsledky budou diskutovány a bude vybráno nejlepší řešení.

7.5 Popis jednotlivých kroků metodiky

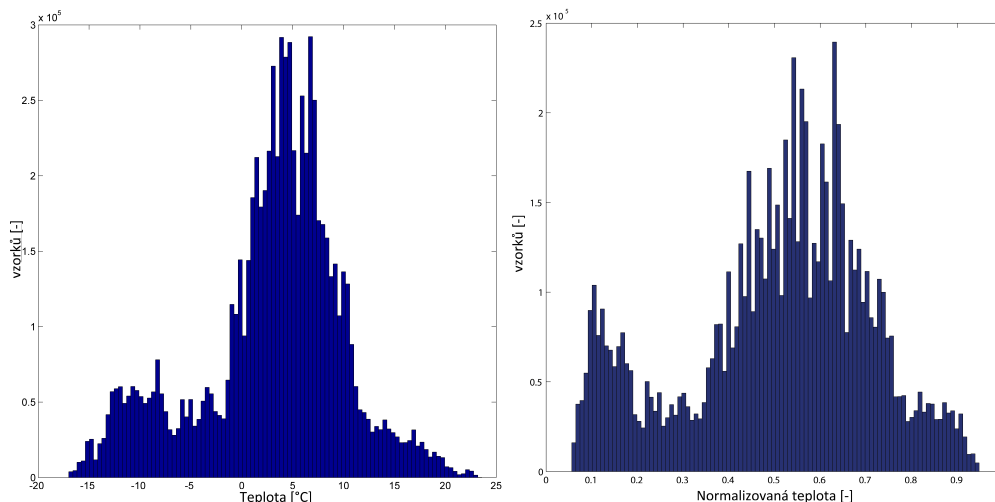
7.5.1 Příprava vstupních dat

Pro vývoj neuronového modelu byly k dispozici dvě množiny vstupních dat. Jeden naměřený vzorek odpovídal jedné sekundě. První množina byla z testovacího provozu, kdy byl zachycen přibližně čtyřdenní cyklus. Tato množina byla použita pouze pro první odhad velikosti neuronové sítě, která bude reagovat na naměřená data. Jejich normalizace do intervalu $\langle 0; 1 \rangle$ proběhla za použití minmax normalizace:

$$f(x) = \frac{x - MIN}{MAX - MIN}, \quad (7.1)$$

kde MIN je minimum ze všech naměřených hodnot dané veličiny x a MAX je maximum ze všech naměřených hodnot dané veličiny.

Druhá, důležitější množina, byla naměřena v průběhu tří měsíců roku 2012. Další data nebyla v době vývoje modelu k dispozici kvůli technické závadě na FSO spoji. Celá množina se tedy sestává z cca osmi milionů hodnot odpovídajícím každé sekundě ve sledovaných třech měsících. V této množině jsou všechna meteorologická data a přijaté úrovně pro oba FSO spoje. Všechna data musí být normalizována do intervalu $\langle 0; 1 \rangle$, protože budeme používat neuronové sítě s aktivními funkcemi s unipolární sigmoidou, konkrétně funkcí Matlabu `logsig()`. Normalizace bude provedena za pomoci logistické funkce (7.2), která je schopna více zvýraznit hodnoty více vzdálené od střední hodnoty. Byl totiž učiněn předpoklad, že přesnost modelu budou více ovlivňovat hodnoty více odlišné od střední hodnoty (například je důležitější, že se teplota změnila o několik stupňů, než že se



Obr. 7.7: Ukázka histogramu teploty před provedením normalizace Obr. 7.8: Histogram teploty po normalizaci funkcí (7.2)

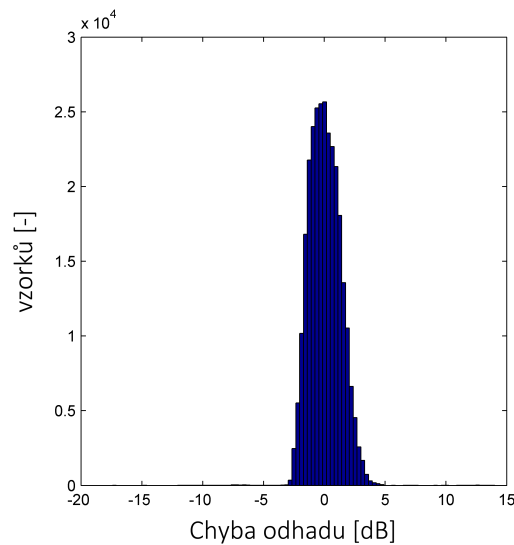
pohnula o jednu desetinu stupně). Logistická normalizační funkce je parametrizována parametry σ a μ , které reprezentují standardní odchylku a střední hodnotu dané změřené veličiny.

$$f(x|\mu, \sigma) = \frac{e^{\frac{x-\mu}{\sigma}}}{\sigma \left(1 + e^{\frac{x-\mu}{\sigma}}\right)^2}, \quad x \in \mathbb{R} \quad (7.2)$$

Efekt použití této funkce na teplotní data je zřejmý z rozdílu histogramu celé množiny naměřených teplotních dat před normalizací - obr. 7.7 - a po provedení normalizace - obr. 7.8. Průběh histogramu v oblasti střední hodnoty se stal plošším a na krajích intervalu se zvýšil počet vzorků. Je samozřejmé, že po vybavení dat neuronovou sítí budou muset být data denormalizována inverzní funkcí se stejnými parametry σ a μ , které měla učící množina.

7.5.2 Výběr architektury a velikosti neuronové sítě

Aby byl k dispozici alespoň hrubý odhad velikosti neuronové sítě, který bude použit pro další zpřesnění modelu, bylo provedeno učení starších dat zbavených nepřesností způsobených výpadky některých čidel [113]. Těchto vzorků bylo přibližně 300000 (stejná množina vzorů, která byla použita pro analýzu shluků pomocí SOM) a z nich bylo uniformně vybrána učící množina o 1000 vzorcích náhodně vybraných z celého průběhu. Byly zkonstruovány dopředné neuronové sítě o velikostech 10x7, 15x10, 25x15, 35x20 a 45x30. Aktivační funkce byla kvůli povaze dat bipolární sigmoida. Tyto velikosti byly vybrány na základě zkušeností s daty v předchozí kapitole. Výhodnocení učení těchto sítí je uvedeno v tab. 7.1. V tabulce je vidět, že nejlepší odhad sítě nastal pro rozměr 25x15, ale i ostatní sítě měly velmi dobrou odezvu a chyba odhadu větší než 5dB nastala v méně než



Obr. 7.9: Histogram chyby pro síť 25x15

0,2% případů. Zde je třeba poznamenat, že se jednalo o relativně krátký časový úsek, ze kterého pocházely vzorky. Histogram průběhu chyby odhadu této sítě je zobrazen na obr. 7.9. Ostatní histogramy mají obdobný průběh a lze je najít v [113].

Síť	10x7	15x10	25x15	35x20	45x30
MSE	0,00355	0,00403	0,00358	0,00327	0,00302
Chyb > 5dB	530	475	366	533	575
	0,19%	0,17%	0,13%	0,19%	0,20%
Chyb > 10dB	168	226	102	102	26
	0,06%	0,08%	0,04%	0,04%	0,01%

Tab. 7.1: Vyhodnocení prvního odhadu rozměru neuronové sítě pro model FSO

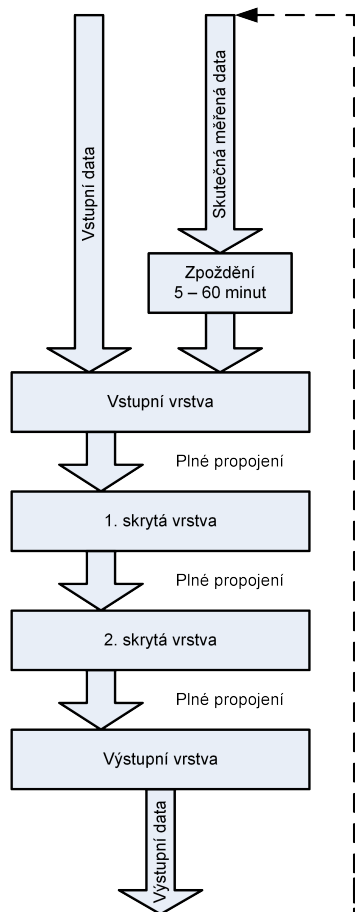
Chyba, kterou jsme dosáhli použitím sítě 25x15, není pro odhad chování FSO spoje dostatečná. Budeme se tedy snažit najít vhodnější uspořádání sítě. Vzhledem k tomu, že stav počasí se mění spojitě v prostoru i čase. Bylo by tedy vhodné do neuronové sítě vnést časový kontext měřených dat. Inspirací může být predikce časových řad. Pro otestování tohoto přístupu byly zvoleny dvě architektury neuronové sítě, které jsou zobrazeny na obr. 7.10, 7.11, a běžná neuronová síť se dvěma skrytými vrstvami. Všechny tyto architektury byly vytvořeny ve třech variantách a to v rozměrech skrytých vrstev 9x6, 14x9 a 17x11. Velikosti těchto sítí byly zvoleny s ohledem na Baum-Hausslerovo pravidlo, síť 17x11 už toto pravidlo výrazně převyšuje.

Neuronová síť na obr. 7.10 má na vstupu zpožděná data útlumů. Jsou vyhodnoceny různé délky zpoždění. Tato architektura neuronové sítě se už blíží k architektuře neuronových sítí a lze říci, že se jedná o speciální případ Nerrandovy

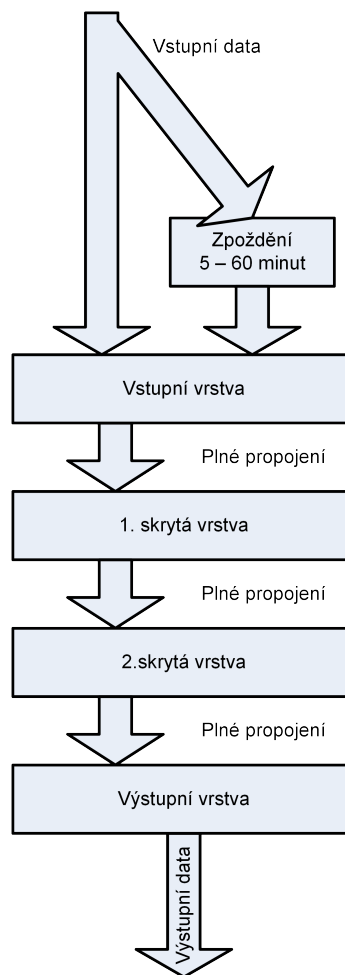
formy neuronové sítě, případně rekurentní neuronové sítě NARX (Nonlinear Autoregressive Network with Exogeneous Inputs) popsané například v [114]. V naší použité variantě byla zpětná vazba realizována skutečnými naměřenými daty.

Neuronová síť na obr. 7.11 není rekurentní. Má dvojnásobnou velikost vstupní vrstvy, kdy druhou polovinu tvoří o pevný časový úsek zpožděná meteorologická data.

Všechny sítě popsané v tomto odstavci jsou plně propojené, mají vstupní vrstvu, která neprovádí žádné váhování vstupů, ale pouze distribuuje vstupní data do první skryté vrstvy. Aktivační funkce všech neuronů ve skrytých vrstvách je unipolární sigmoida (funkce $\text{logsig}()$), všechny neurony mají proměnnou prahovou hodnotu. Výstupní vrstva obsahuje dva neurony s aktivační funkcí $\text{logsig}()$. Každý neuron se stará o predikci úrovně jednoho z FSO spojů.



Obr. 7.10: Architektura neuronové sítě se zpožděnými daty útlumů



Obr. 7.11: Architektura neuronové sítě se zpožděnými meteorologickými daty

7.5.3 Konstrukce učících množin

Po normalizaci byly z celého souboru dat vybrány čtyřhodinové bloky a každý třetí byl vložen do trénovací množiny. Bylo tak zajištěno, že se v trénovací množině objeví data ze dne i z noci, vzhledem k tomu, že FSO spoje vykazují odchylky i na základě střídání dne a noci. U vytvořené učící množiny bylo vždy zkontrolováno podle histogramu, jestli učící množina zhruba odpovídá všem naměřeným datům.

7.5.4 Učení neuronové sítě

Celkem bylo tedy vytvořeno 9 neuronových sítí, které byly celkem 27krát učeny. Toto učení muselo proběhnout vícekrát (v tomto případě třikrát) a byl vybrán vždy nejlepší výsledek.

Učení probíhalo za použití GDM backpropagation [115]. Tento algoritmus je považován za odolnější než základní backpropagation. Trénování každé sítě probíhalo vždy alespoň 5000 epoch a učící konstanta se postupně snižovala z počáteční hodnoty 0.3 až k 0.05. Konstanta ovlivňující setrvačnost algoritmu klesala z 0,9 až k hodnotě 0,6. Jako validační množina byla v tomto případě použita množina o velikosti 15% všech naměřených hodnot. Touto množinou bylo ověřováno, zda nedochází k přeučení neuronové sítě (postup popsáný v [116]). Učení všech sítí tak bylo zastaveno dříve, než by mohlo dojít k overfittingu, tj. v okamžiku, kdy MSE validační množiny začíná růst.

7.5.5 Vyhodnocení jednotlivých neuronových sítí

Tabulka 7.2 ukazuje vyhodnocení chyby pro nejlépe natrénovanou síť 14x9 pro spoj WaveBridge. Z tabulek 7.3 a 7.4 lze rozpoznat velmi malé odchylky od nejlepší sítě. V těchto tabulkách byly pro lepší přehlednost vynechány chyby, které jsou menší než 0.7 dB, kterou můžeme považovat za malou. Tabulka 7.5 prezentuje nejlepší nalezenou odezvu - naučenou síť 17x11. Mezi dalšími velikostmi sítí však nebyly znatelné rozdíly v počtech chyb (rozdíly ve velikostech jednotlivých tříd chyb nebyly větší než desítky případů). Je třeba znovu připomenout, že jedna chyba odpovídá jedné sekundě z naměřených tří měsíců.

K pochopení tabulek je třeba znát význam jednotlivých konfigurací:

- C1 - na vstupu sítě jsou pouze meteorologická data
- C2 - meteorologická data + 5 min zpožděná meteorologická data
- C3 - meteorologická data + 15 min zpožděná meteorologická data
- C4 - meteorologická data + 30 min zpožděná meteorologická data
- C5 - meteorologická data + 60 min zpožděná meteorologická data
- C6 - meteorologická data + 5 min zpožděný útlum spojů
- C7 - meteorologická data + 5 min zpožděný útlum spojů
- C8 - meteorologická data + 5 min zpožděný útlum spojů
- C9 - meteorologická data + 5 min zpožděný útlum spojů

Zabývejme se nejdříve spojem WaveBridge. Pro tento spoj lze najít nejlepší konfigurace C6-C9. To znamená, že navržená rekurentní síť je nejlepším zatím dostupným prostředkem pro modelování vlivu počasí na FSO. Je ale vidět i značné zlepšení v modelu, který obsahuje zpožděná meteorologická data. Zřetelně nejhorší odezvu má neuronová síť bez jakýchkoli opatření s výjimkou nejmenší sítě 9x6, jejíž odezva se může rovnat s většími sítěmi, které mají na vstupech zpožděná meteorologická data. Síť 14x9 poskytla v konfiguraci C6 nejlepší odezvu - 501 chyb větších než 0.7 dB. To odpovídá přibližně 8 minutám a 20 sekundám ve třech měsících. To je již velmi dobrý výsledek. Rozložení chyby tohoto odhadu v čase je vykreslen na obr. 7.12. Je vidět, že velké chyby se objevují osamoceně, proto lze usuzovat, že v tu chvíli nastal efekt, který nemusí nutně souviset s počasím, ale může se jednat třeba o přelet ptáka, průchod člověka a podobně. Případně se může jednat o zvláštní jev, který nebyl v učící množině. Při analýze meteorologických údajů v čase, kdy nastaly chyby v odhadu, jsem narazil na fakt, že chyby nenastávaly při bezvětří, ale spíše při větru od 2 m/s výš (na obr. 7.13 je zobrazena kombinace teploty a rychlosti větru pro okamžiky, kdy model neposkytl správnou odezvu). Při analýze chyb byly vysledovány následující skutečnosti o meteorologických datech:

- Přibližně 50% všech chyb nastalo při minutovém průměru větru 5 m/s nebo nárazech 6-8 m/s.
- Nejvíce chyb nastalo při velmi lehkém dešti (0.1-0.2 mm/min).
- Nejvíce chyb nastalo při relativní vlhkosti 81-83% (téměř polovina všech chyb).
- V 50% chybných případech vál severní vítr.
- Teplota neměla na chyby výraznější vliv. Průměrná teplota byla 3,5°C, což odpovídá střední hodnotě za dané období.
- Ve 2/3 případů byl atmosferický tlak 990 hPa.

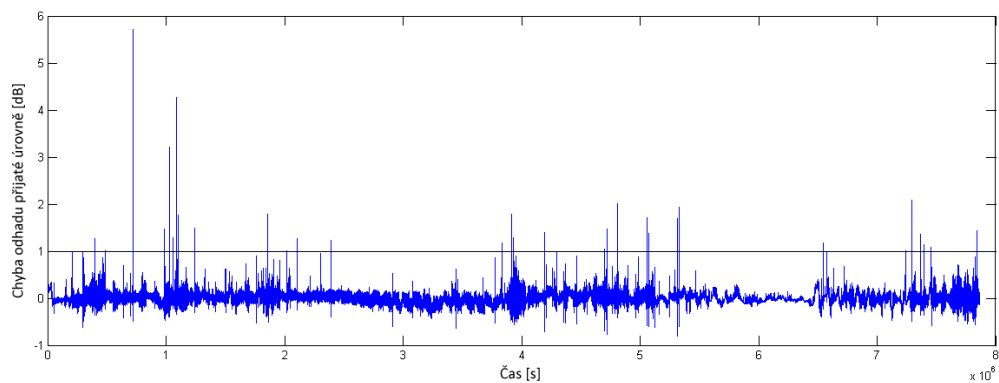
Nyní se zaměříme na odhad spoje FlightStrata. Z tohoto spoje máme informaci pouze ve formě deseti tříd, kdy každá třída odpovídá poklesu o 6 dB. Pro tato data se evidentně dostáváme do problému s granularitou. Model pro tento spoj chybuje v 6-7% případů. Pro všechny sítě je tato chyba téměř stejná. Lze tedy tvrdit, že tento spoj je pro požadované modelování nevhodné a bylo by potřeba zjistit jemnější informaci o přijaté úrovni signálu.

Sít' 14x9 Konfigurace	Velikost chyby [dB]								Suma
	-21.0	-7.0	-4.2	-2.2	-1.5	(± 0.8)	+1	+1.5	
C1	0	2	6	573	641	8040	447	0	9709
C2	0	2	5	697	323	52	0	0	1079
C3	0	1	6	554	305	2248	1299	0	4413
C4	0	2	6	566	386	3480	350	0	4790
C5	0	1	7	623	518	2153	96	0	3398
C6	0	2	12	366	119	1	1	0	501
C7	0	1	10	456	183	0	0	0	650
C8	0	1	14	545	236	0	0	0	796
C9	0	1	14	553	292	0	0	0	860

Tab. 7.2: Vyhodnocení chyby pro síť o velikosti 14x9 pro odhad chování spoje WaveBridge

Sít' 17x11 Konfigurace	Velikost chyby [dB]								Suma
	-21.0	-7.0	-4.2	-2.2	-1.5	(± 0.8)	+1	+1.5	
1	0	1	7	526	427	2437	1	0	3399
2	0	2	5	624	680	1805	255	0	3371
3	0	2	5	571	423	2	0	0	1003
4	0	2	8	729	350	3199	93	0	4381
5	0	2	6	507	344	1390	1	0	2250
6	0	1	11	463	171	0	0	0	646
7	0	1	14	461	173	0	0	0	649
8	0	1	12	611	250	0	0	0	874
9	0	2	12	701	530	0	0	0	1245

Tab. 7.3: Vyhodnocení chyby pro síť o velikosti 17x11 pro odhad chování spoje WaveBridge



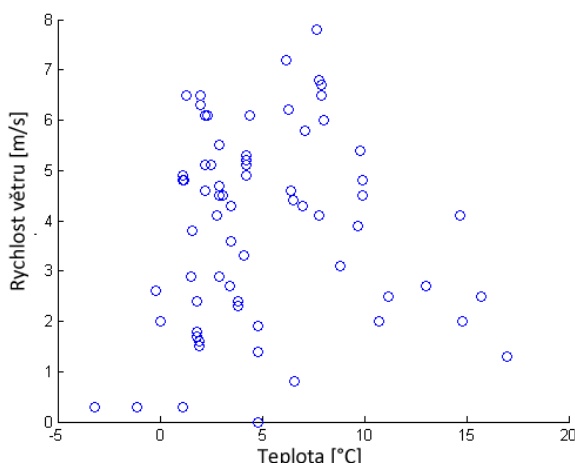
Obr. 7.12: Časový průběh velikosti chyby odhadu pro neuronovou síť 14x9 v konfiguraci C6

Sít 9x6 Konfigurace	Velikost chyby [dB]								Suma
	-21.0	-7.0	-4.2	-2.2	-1.5	(±0.8)	+1	+1.5	
1	0	2	6	611	307	2	0	0	928
2	0	2	5	578	469	611	1	0	1666
3	0	2	5	460	328	2023	48	0	2866
4	0	2	5	595	390	2815	12	0	3819
5	0	2	4	481	359	7	0	0	853
6	0	2	12	450	165	1	0	0	630
7	0	2	18	529	183	1	0	0	733
8	0	1	10	581	403	0	0	0	995
9	0	2	18	652	394	0	0	0	1066

Tab. 7.4: Vyhodnocení chyby pro síť o velikosti 9x6 pro odhad chování spoje WaveBridge

Sít 17x11 Konf.	Velikost chyby [dB]							Suma
	-24	-21	-12	-6	0	6	12	
1	0	22408	39106	357416	7369656	64850	11027	494807
2	0	22408	39109	357443	7369399	64777	11027	494764
3	0	21939	38491	355851	7354742	81813	11027	509121
4	0	22408	39109	357129	7363552	69438	11027	499111
5	0	22408	39109	357427	7365706	65186	11027	495157
6	0	22408	39109	357443	7369399	64777	11027	494764
7	13	22424	39205	358792	7367926	64812	10691	495937
8	0	22061	38539	350346	7363529	77159	11029	499134
9	5	22352	36235	350409	7367653	73165	11044	493210

Tab. 7.5: Vyhodnocení chyby pro nejlépe natrénovanou síť 17x11 pro spoj Flight-Strata



Obr. 7.13: Průměrná rychlost větru v závislosti na teplotě pro simulační chyby větší než 1 dB.

7.6 Dílčí závěr

V této kapitole byla představena funkční metodika pro tvorbu modelu bezdrátového optického spoje pro predikci jeho chování v závislosti na aktuálních povětrnostních podmínkách. Byly popsány jednotlivé kroky metodiky od tvorby učicích a validačních množin až ke způsobu vyhodnocování odezvy modelu. Zároveň bylo zjištěno, že pro kvalitní model spoje je nutné mít dostatečně jemná data o přijaté úrovni signálu na jednotlivých FSO hlavicích.

Byla provedena analýza chybných odezev a bylo zjištěno, že by bylo vhodné se při dalším zpřesňování modelu zaměřit na okamžiky, kdy je mírný déšť a vítr nad 5 m/s. V případě deště se může jednat o chybu způsobenou měřením, protože použitý srážkoměr nemusí pro tak malá množství srážek poskytovat zcela přesné informace.

7.7 Doporučení pro další zlepšení

Pro lepší ověření kvality modelu je třeba získat větší množství dat - ideálně z několika let. Současný model je platný pouze pro malou část roku ve které byla data měřena.

Stále existuje potenciál pro zlepšení odezvy vytvořeného modelu:

- Je možnost zakomponovat více historických kroků do rekurentního neurového modelu.
- Bylo by vhodné využít ke zpřesnění modelu i úroveň přijatého signálu na druhé straně spoje, která v tuto chvíli není měřena
- Model může být doplněn o kontinuální učení.

- Lze použít další druhy neuronových sítí a ověřit jejich efektivitu.

Bylo by vhodné zajistit kontinuální měření průběhu přijaté úrovně, aby bylo možné sledovat i rychlé turbulence. Odhad chování turbulencí je relativně otevřenou oblastí, která by si zasloužila další výzkum.

Kapitola 8

Závěr

V práci byly ověřeny možnosti využití dat experimentálního mikrovlnného spoje v kombinaci s umělými neuronovými sítěmi pro klasifikaci množství deště v jeho průběhu. Vzhledem k povaze sledovaného jevu a možnostem experimentálního měřicího spoje, který ale nebyl cíleně určen k řešení daného problému, se podařilo vytvořit použitelný model odhadu s přesností podrobněji popsanou v textu a dílčích závěrech příslušných kapitol.

Poznatky a postupy z klasifikace deště byly dále využity při sestavení metodiky využití neuronových sítí pro radiový odhad gradientu refraktivity a výškového profilu gradientu refraktivity na spoji Praha-Poděbrady. Přesnost odhadu korespondovala předpokládané chybě, která byla způsobena způsobem měření meteorologických dat na experimentálním spoji. Samotné fyzikální principy, které se uvnitř neuronového modelu skrývají, nebyly předmětem této práce.

V další části práce věnované modelování atmosferických vlivů na FSO za použití neuronových sítí a běžně dostupných meteorologických senzorů byla navržena metodika, která poskytuje velmi dobré přesnosti odhadu. Zároveň byly rozebrány chyby odhadu, které model nebyl schopen správně predikovat. Tyto chyby mohly být způsobeny způsobem měření meteorologických veličin (především deště).

Cíle práce vytyčené v úvodu práce se podařilo splnit. Samotnou metodiku pro práci s daty podobné povahy lze doporučit především pro krátké spoje, kde nedochází k neměřitelným výkyvům atmosferických vlivů. V případě, kdy by informace o těchto vlivech byly dostupné, lze je shodným přístupem do metodiky doplnit.

Hlavním přínosem této práce je samotná metodika pro zpracování meteorologických dat v kombinaci s měřenými parametry komunikačních spojů, která zatím nebyla jinde publikována. Využití postupů popsanych v této práci lze mimo jiné najít i v dalším zamýšleném rozvoji problematiky a tím je využití dat z komunikačních spojů k diagnostice aktuálního stavu atmosféry - tedy k opačnému přístupu, než jaký byl popsán v této práci.

Literatura

- [1] Rui Xu and II Wunsch, D. Survey of clustering algorithms. *Neural Networks, IEEE Transactions on*, 16(3):645–678, May 2005.
- [2] Dong Huang, Jian-Huang Lai, and Chang-Dong Wang. Incremental support vector clustering with outlier detection. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 2339–2342, Nov 2012.
- [3] Li-Ying Wu and Jeen-Shing Wang. A simplified support vector clustering algorithm. In *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on*, pages 1259–1264, Oct 2008.
- [4] Sandhya Samarasinghe. *Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition*. Auerbach Publisher, 1 edition, September 2006. ISBN: 978-0849333750.
- [5] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and Regression Trees*. Wadsworth Statistics/Probability. Chapman and Hall/CRC, 1 edition, January 1984. ISBN: 978-0412048418, Kindle edition.
- [6] Koen-Myung Lee, Kyung-Mi Lee, Jee-Hyong Lee, and Hyung Lee-Kwang. A fuzzy decision tree induction method for fuzzy data. In *Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE '99. 1999 IEEE International*, volume 1, pages 16–21 vol.1, Aug 1999.
- [7] K. Hornik, Stinchcombe M., and H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.
- [8] D. M. Bates and D. G. Watts. *Nonlinear Regression Analysis and Its Applications*. Wiley-Interscience, 1 edition, April 2007. ISBN: 978-0470139004.
- [9] J. D. Hamilton. *Time Series Analysis*. Princeton University Press, 1 edition, 1994. ISBN: 9780691042893.
- [10] W. Hardle. *Applied Nonparametric Regression*. Number 352 in Econometric Society Monographs. Cambridge University Press, 1 edition, January 1992. ISBN: 978-0521429504.

- [11] ITU-R P.530 - Propagation Data and Prediction Methods Required for the Design of Terrestrial Line-of-sight Systems, 2013.
- [12] ITU-R P.676 - Attenuation by Atmospheric Gases, 2013.
- [13] ITU-R P.836 - Water vapour: surface density and total columnar content, 2013.
- [14] ITU-R P.840 - Attenuation due to Clouds and Fog, 2013.
- [15] ITU-R P.837 - Characteristics of Precipitation for Propagation Modelling, 2012.
- [16] ITU-R P.838 - Specific Attenuation Model for Rain for Use in Prediction Methods, 2005.
- [17] O. Ozgun, M. Kuzuoglu, G. Apaydin, and L. Sevgi. Two-way split-step parabolic equation algorithm for tropospheric propagation: Tests and comparisons. In *Microwave Symposium (MMS), 2010 Mediterranean*, pages 14–17, Aug 2010.
- [18] C.L. Rino and V.R. Kruger. A comparison of forward-boundary-integral and parabolic-wave-equation propagation models. *Antennas and Propagation, IEEE Transactions on*, 49(4):574–582, Apr 2001.
- [19] G.D. Dockery. Modeling electromagnetic wave propagation in the troposphere using the parabolic equation. *Antennas and Propagation, IEEE Transactions on*, 36(10):1464–1470, Oct 1988.
- [20] J.H. Richter. Sensing of radio refractivity and aerosol extinction. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 1, pages 381–385 vol.1, Aug 1994.
- [21] C.H. Wash and K.L. Davidson. Remote measurements and coastal atmospheric refraction. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International*, volume 1, pages 397–401 vol.1, Aug 1994.
- [22] F. Cuccoli and L. Facheris. Rainfall monitoring based on microwave link attenuation and related tomographic processing. In *Advances in Radar and Remote Sensing (TyWRRS), 2012 Tyrrhenian Workshop on*, pages 172–176, Sept 2012.
- [23] P. Valtr, P. Pechac, V. Kvicera, and M. Grabner. Estimation of the refractivity structure of the lower troposphere from measurements on a terrestrial

- multiple-receiver radio link. *Antennas and Propagation, IEEE Transactions on*, 59(5):1707–1715, May 2011.
- [24] K. Allen, R. Ott, E.J. Violette, and R.H. Espeland. Height-gain study for 23 km links at 9.6, 11.4, and 28.8 GHz. *Antennas and Propagation, IEEE Transactions on*, 30(4):734–740, Jul 1982.
- [25] W. Lam and Alan R. Webster. Microwave propagation on two line-of-sight oversea paths. *Antennas and Propagation, IEEE Transactions on*, 33(5):510–516, May 1985.
- [26] O. Sasaki and T. Akiyama. Studies on radio duct occurrence and properties. *Antennas and Propagation, IEEE Transactions on*, 30(5):853–858, Sep 1982.
- [27] P. Gerstoft, D.F. Gingras, L.T. Rogers, and W.S. Hodgkiss. Estimation of radio refractivity structure using matched-field array processing. *Antennas and Propagation, IEEE Transactions on*, 48(3):345–356, Mar 2000.
- [28] G.S. Woods, D.L. Maskell, and A.J. Kerans. New angle-of-arrival measurement technique for over ocean propagation studies. In *Communications Systems, 2004. ICCS 2004. The Ninth International Conference on*, pages 616–620, Sept 2004.
- [29] P. Valtr and P. Pechac. The influence of horizontally variable refractive index height profile on radio horizon range. *Antennas and Wireless Propagation Letters, IEEE*, 4(1):489–491, 2005.
- [30] J. Agnew. Lidar and radar tropospheric profiling at chibolton observatory. In *Proceedings of 6th International Symposium on Tropospheric Profiling*, pages 151–153, Leipzig, 2003.
- [31] S. Hosseinzadeh and N. Samsunchi. The troposphere refractivity slop determination from propagation loss by the artificial neural networks. In *Telecommunications, 2008. IST 2008. International Symposium on*, pages 88–91, Aug 2008.
- [32] P. Basili, S. Bonafoni, V. Mattioli, F. Pelliccia, and P. Ciotti. Neural networks for tropospheric profiling from GPS-LEO radio occultation. In *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, pages 1693–1696, July 2007.
- [33] O. Jicha, P. Pechac, V. Kvicera, and M. Grabner. Estimation of the radio refractivity gradient from diffraction loss measurements. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(1):12–18, Jan 2013.

- [34] J. S. Marshall and W. M. K. Palmer. The distribution of raindrops with size. *Journal of Meteorology*, 5:165–166, 1948.
- [35] J. S. Marshall, W. Hitschfeld, and K. L. S. Gunn. Advances in radar weather. *Advances in Geophysics*, 5(2), 1955.
- [36] P. Dvorak. *Atlas oblaku*. Svet kridel, Cheb, 2012.
- [37] COST 253 Committee. *Radiowave propagation effects on next-generation fixed-services terrestrial telecommunication systems*. European Commission, 1996. ISBN: 92-827-8023-6, EUR: 16992.
- [38] G. Lefrancois. Modele theoretique de precipitation equivalents sur un trajet radioelectrique. *Ann. Telecom*, 26(11-12):795–800, 1971.
- [39] W. L. Stutzman and W. K. Dishman. A simple model for the estimation of rain induced attenuation along earth-spacepaths at millimeter wavelengths. *Radio Science*, 17(6):1465–1476, 1982.
- [40] S. H. Lin. Nationwide long-term rain statistics and empirical calculation of 11 GHz microwave rain attenuation. *Bell System Technical Journal*, 56(9):1581–1604, 1977.
- [41] V. I. Tatarski. *The Effects of the Turbulent Atmosphere on Wave Propagation*. McGraw-Hill, New York, 1971.
- [42] P. Pechač and S. Zvánovec. *Zaklady sireni vln pro planovani radiovych spoju*. BEN, 1 edition, 2007. ISBN: 978-80-7300-223-7.
- [43] ITU-R P.453 -The Radio Refractive Index: Its Formula and Refractivity Data, 2012.
- [44] P. Valtr and P. Pechac. Diffraction calculations and measurements in millimeter frequency band. *Radioengineering*, 13(3):18–21, September 2004.
- [45] R. Akbarpour and Alan R. Webster. Ray-tracing and parabolic equation methods in the modeling of a tropospheric microwave link. *Antennas and Propagation, IEEE Transactions on*, 53(11):3785–3791, Nov 2005.
- [46] Isaac I. Kim, Bruce McArthur, and Eric J. Korevaar. Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications. *Proc. SPIE*, 4214:26–37, 2001.
- [47] P. W. Kruse, L. D. McGlauchlin, and R. B. McQuistan. *Elements of infrared technology: generation, transmission, and detection*. John Wiley and Sons, 1962. ISBN: 978-0471508861.

- [48] F. Nadeem, T. Javornik, E. Leitgeb, V. Kvicera, and G. Kandus. Continental fog attenuation empirical relationship from measured visibility data. *Radioengineering*, 19(4):596–600, 2010.
- [49] L. C. Andrews and R. L. Phillips. *Laser Beam Propagation through Random Media*. SPIE Press, September 2005.
- [50] Xiang Yi, Zengji Liu, and Peng Yue. Optical scintillations and fade statistics for FSO communications through moderate-to-strong non-kolmogorov turbulence. *Optics & Laser Technology*, 47(0):199 – 207, 2013.
- [51] J. Libich and S. Zvanovec. Influence of turbulence in near vicinity of buildings on free-space optical links. *IET Microwaves, Antennas and Propagation*, 5(9):1039–1044, 2011.
- [52] P. Barcik, L. Hudcova, and O. Wilfert. Influence of the atmospheric turbulence on the laser beam. In *Network and Optical Communications (NOC), 2013 18th European Conference on and Optical Cabling and Infrastructure (OC i), 2013 8th Conference on*, pages 83–86, July 2013.
- [53] M. Loschnigg, P. Mandl, and E. Leitgeb. Long-term performance observation of a free space optics link. In *Telecommunications, 2009. ConTEL 2009. 10th International Conference on*, pages 305–310, June 2009.
- [54] M. Loeschnigg, T. Plank, and E. Leitgeb. Five years analysis of a free space optics link in Graz. In *Antennas and Propagation (EuCAP), 2012 6th European Conference on*, pages 1248–1251, March 2012.
- [55] M. Tatarko, L. Ovsenik, and J. Turan. Properties of hybrid FSO/RF link with 60 GHz RF backup link. In *Information Communication Technology Electronics Microelectronics (MIPRO), 2013 36th International Convention on*, pages 495–497, May 2013.
- [56] J. Libich, M. Mudroch, P. Dvorak, and S. Zvanovec. Performance analysis of hybrid FSO/RF link. In *EuCAP 2012*, pages 1235–1238, March 2012.
- [57] Z. Kolka, Z. Kincl, V. Biolkova, and D. Biolek. Hybrid FSO/RF test link. In *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2012 4th International Congress on*, pages 502–505, Oct 2012.
- [58] F. Nadeem, E. Leitgeb, M.S. Awan, and G. Kandus. Comparing the snow effects on hybrid network using optical wireless and GHz links. In *Satellite and Space Communications, 2009. IWSSC 2009. International Workshop on*, pages 171–175, Sept 2009.
- [59] F. Nadeem, E. Leitgeb, O. Koudelka, T. Javornic, and G. Kandus. Comparing the rain effects on hybrid network using optical wireless and GHz links.

- In *Emerging Technologies, 2008. ICET 2008. 4th International Conference on*, pages 156–161, Oct 2008.
- [60] O. Bouchet. *Free-space optics: propagation and communication*. Wiley-ISTE, January 2006.
- [61] I.I. Kim et al. Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications. In *Proc. SPIE 4214*, 2001.
- [62] M. A. Naboulsi and et al. Fog attenuation prediction for optical and infrared waves. *Journal of Optical Engineering*, 43:319–329, 2004.
- [63] M. Grabner and V. Kvicera. The wavelength dependent model of extinction in fog and haze for free space optical communication. *Journal of Optics Express*, 19:3379–3386, 2012.
- [64] M. Ijaz, Z. Ghassemlooy, J. Perez, V. Brazda, and O. Fiser. Enhancing the atmospheric visibility and fog attenuation using a controlled FSO channel. *IEEE Photonics Technology Letters*, 25(13):1262–1265, 2013.
- [65] A. Kolmogorov. *Classic Papers on Statistical Theory*. Wiley-Interscience, 1961.
- [66] S. Zvanovec, J. Perez, Z. Ghassemlooy, S. Rajbhandari, and J. Libich. Route diversity analyses for free-space optical wireless links within turbulent scenarios. *Journal of Optics Express*, 21:7641–7650, 2013.
- [67] R. M. Goody and Y. L. Yung. *Atmospheric Radiation: Theoretical Basis*. Oxford University Press, 2nd edition, 1989.
- [68] K. T. Whitby and B. Cantrell. Atmospheric aerosols - characteristics and measurements. In *International Conference on Environmental Sensing and Assessment, Las Vegas, Nevada*. IEEE, New York, 1976.
- [69] H. Moosmuller, R.K. Chakrabarty, and W.P. Arnott. Aerosol light absorption and its measurement: A review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(11):844 – 878, 2009. Light Scattering: Mie and More Commemorating 100 years of Mie’s 1908 publication.
- [70] D. W. Hahn. Light scattering theory. Technical report, Department of Mechanical and Aerospace Engineering, University of Florida, July 2009.
- [71] T. Wriedt. *The Mie Theory - Basics and Applications*, volume 169 of *Springer Series in Optical Sciences*, chapter Mie Theory: A Review, page 259. Springer, 2012. ISBN 978-3-642-28737-4.

- [72] Michael I. Mishchenko. Electromagnetic scattering by nonspherical particles: A tutorial review. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(11):808 – 832, 2009. Light Scattering: Mie and More Commemorating 100 years of Mie’s 1908 publication.
- [73] I. I. Kim, M. Mitchell, and E. J. Korevaar. Measurement of scintillation for free-space laser communication at 785 nm and 1550 nm. *Proc. SPIE*, 3850:49–62, 1999.
- [74] M. Gebhart, E. Leitgeb, and J. Bregenzer. Atmospheric effects on optical wireless links. In *Telecommunications, 2003. ConTEL 2003. Proceedings of the 7th International Conference on*, volume 2, pages 395–401 vol.2, June 2003.
- [75] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408, 1958.
- [76] F. Rosenblatt. Perceptron simulation experiments. *Proceedings of the IRE*, 48(3):301–309, 1960.
- [77] F. Rosenblatt. Analytic techniques for the study of neural nets. *Applications and Industry, IEEE Transactions on*, 83(74):285–292, 1964.
- [78] B. Widrow. Pattern recognition and adaptive control. *Applications and Industry, IEEE Transactions on*, 83(74):269–277, Sept 1964.
- [79] B. Widrow, R.G. Winter, and R.A. Baxter. Layered neural nets for pattern recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(7):1109–1118, Jul 1988.
- [80] M. L. Minsky and S. A. Papert. *Perceptrons*. MIT Press, Cambridge, MA, USA, 1969.
- [81] R.P. Lippmann. An introduction to computing with neural nets. *ASSP Magazine, IEEE*, 4(2):4–22, Apr 1987.
- [82] R. O. Winder. Single stage threshold logic. In *Switching Circuit Theory and Logical Design, 1961. SWCT 1961. Proceedings of the Second Annual Symposium on*, pages 321–332, Oct 1961.
- [83] Philip D. Wasserman. *Neural computing: Theory and Practice*. Van Nostrand Reinhold, 1989.
- [84] R. Rojas. *Neural Networks: A Systematic Introduction*. Springer Berlin Heidelberg, 1996. ISBN 978-3-642-61068-4.

- [85] A. J. Mansfield. Comparison of perceptron training by linear programming and by the perceptron convergence procedure. In *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, volume ii, pages 25–30 vol.2, Jul 1991.
- [86] Eric B. Baum. Polynomial time algorithms for learning neural nets. In *Proceedings of the Third Annual Workshop on Computational Learning Theory, COLT '90*, pages 258–272, San Francisco, CA, USA, 1990. Morgan Kaufmann Publishers Inc.
- [87] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chapter Learning Internal Representations by Error Propagation, pages 318–362. MIT Press, Cambridge, MA, USA, 1986.
- [88] B. Widrow and M. E. Hoff. Adaptive switching circuits. In *1960 IRE WESCON Convention Record*, volume 4, pages 96–104, New York, 1960. IRE.
- [89] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, (323):533–536, 1986.
- [90] T. Samad. Back-propagation is significantly faster if the expected value of the source unit is used for update. Technical report, Honeywell CSDD, 1000 Boone Ave. N., Golden Valley, MN 55427, 1988.
- [91] S. E. Fahlman. Faster-learning variations on back-propagation: An empirical study. In *1988 Connectionist Models Summer School*. Morgan-Kaufmann, Los Altos CA, 1988.
- [92] Christoph von der Malsburg. A model for self-organization of orientation sensitivity and columns in the visual cortex. *Pflügers Arch. Suppl. to*, 339, 1973. R 95.
- [93] Christoph von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:85–100, 1973.
- [94] David J. Willshaw and Christoph von der Malsburg. The establishment of patterned neural connections according to a theory of self - organization. *Pflügers Arch. Suppl. to*, 359:463–469, 1975.
- [95] David J. Willshaw and Christoph von der Malsburg. How patterned neural connections can be set up by self-organization. In *Proceedings of the Royal Society London*, volume B194, pages 431–445, 1976.

- [96] Joseph Sirosh and Risto Miikkulainen. Self-organizing feature maps with lateral connections: Modeling ocular dominance. In M. C. Mozer, P. Smolensky, D. S. Touretzky, J. L. Elman, and A. S. Weigend, editors, *Proceedings of the 1993 Connectionist Models Summer School*, pages 31–38, 1994.
- [97] Melody Y. Kiang. Extending the kohonen self-organizing map networks for clustering analysis. *Computational Statistics & Data Analysis*, (38):161–180, 2001.
- [98] P.J. Werbos. Backpropagation: past and future. In *Neural Networks, 1988., IEEE International Conference on*, pages 343–353 vol.1, July 1988.
- [99] O. Nerrand, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, and S. Marcos. Neural networks and nonlinear adaptive filtering: Unifying concepts and new algorithms. *Neural Computation*, 5(2):165–199, March 1993.
- [100] L. Personnaz and G. Dreyfus. Comment on "discrete-time recurrent neural network architectures: A unifying review". *Neurocomputing*, 20(1-3):325 – 331, 1998.
- [101] Reinhard Teschl, Walter L. Randeu, and Franz Teschl. Improving weather radar estimates of rainfall using feed-forward neural networks. *Neural Networks*, 20(4):519 – 527, 2007. Computational Intelligence in Earth and Environmental Sciences.
- [102] S.P. Sotiroudis, S.K. Goudos, K.A. Gotsis, K. Siakavara, and J.N. Sahalos. Modeling by optimal artificial neural networks the prediction of propagation path loss in urban environments. In *Antennas and Propagation in Wireless Communications (APWC), 2013 IEEE-APS Topical Conference on*, pages 599–602, Sept 2013.
- [103] Shahrokh Asadi, Jamal Shahrabi, Peyman Abbaszadeh, and Shabnam Tabanmehr. A new hybrid artificial neural networks for rainfall and runoff process modeling. *Neurocomputing*, 121(0):470 – 480, 2013. Advances in Artificial Neural Networks and Machine Learning Selected papers from the 2011 International Work Conference on Artificial Neural Networks (IWANN 2011).
- [104] S. Khedairia and M.T. Khadir. Self-organizing map and k-means for meteorological day type identification for the region of annaba, algeria. In *Computer Information Systems and Industrial Management Applications, 2008. CISIM '08. 7th*, pages 91–96, June 2008.
- [105] C.N. Stephanidis, A.P. Cracknell, and L. W B Hayes. The implementation of self organised neural networks for cloud classification in digital satellite images. In *Geoscience and Remote Sensing Symposium, 1995. IGARSS '95*.

- 'Quantitative Remote Sensing for Science and Applications', *International*, volume 1, pages 455–457 vol.1, Jul 1995.
- [106] D.E. Skinner, J.D. Connor, S.Y. Foo, M.H. Weatherspoon, and N. Powell. Optimization of a multi-band reconfigurable microstrip line-fed rectangular patch antenna using self-organizing maps. In *Wireless and Microwave Technology Conference, 2009. WAMICON '09. IEEE 10th Annual*, pages 1–4, April 2009.
- [107] J. Vermaak and E.C. Botha. Recurrent neural networks for short-term load forecasting. *Power Systems, IEEE Transactions on*, 13(1):126–132, Feb 1998.
- [108] T.G. Barbounis, J.B. Theocharis, M.C. Alexiadis, and P.S. Dokopoulos. Long-term wind speed and power forecasting using local recurrent neural network models. *Energy Conversion, IEEE Transactions on*, 21(1):273–284, March 2006.
- [109] Xiang Sun and Yong Ni. Recurrent neural network with kernel feature extraction for stock prices forecasting. In *Computational Intelligence and Security, 2006 International Conference on*, volume 1, pages 903–907, Nov 2006.
- [110] Martin Mudroch, Martin Grabner, Vaclav Kvicera, Premysl Hudec, Petr Bedrna, and Pavel Pechac. An experimental study of the lowest troposphere layers at 10 GHz - first results. In J Sistek, K Novotna, M Pourova, L Cifra, and L Visek, editors, *PROCEEDINGS OF COMITE 2008*, pages 73–76, 2008.
- [111] E.B. Baum and D. Haussler. What size net gives valid generalization? *Neural Computation*, 1(1):151–160, 1989.
- [112] O. Jicha, P. Pechac, V. Kvicera, and M. Grabner. On the uncertainty of refractivity height profile measurements. *Antennas and Wireless Propagation Letters, IEEE*, 10:983–986, 2011.
- [113] M. Mudroch, J. Libich, S. Zvanovec, and M. Mazanek. FSO link performance modelling using artificial intelligence. In *Antennas and Propagation (EUCAP), Proceedings of the 5th European Conference on*, pages 1715–1718, 2011.
- [114] H.T. Siegelmann, B.G. Horne, and C.L. Giles. Computational capabilities of recurrent NARX neural networks. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 27(2):208–215, 1997.
- [115] M.T. Hagan, H.B. Demuth, and M.H. Beale. *Neural Network Design*. PWS Publishing, Boston, MA, 1996.

- [116] Y. Wada and M. Kawato. Estimation of generalization capability by combination of new information criterion and cross validation. In *Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on*, volume ii, pages 1–6 vol.2, 1991.