

دانشگاه صنعتی شاهرود

دانشکده مهندسی معدن و ژئوفیزیک

پایان نامه کارشناسی ارشد ژئوفیزیک - گرایش ژئوالکتریک

مدل سازی معکوس دوبعدی داده های مگنتوتلوریک با استفاده از شبکه های
عصبی مصنوعی

محقق:

فاطمه طهماسبی

اساتید راهنما:

دکتر محمد مهدی فاتح

دکتر علی مرادزاده

استاد مشاور:

مهندس مصطفی علامه زاده

تابستان ۱۳۸۳

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تقدیم بہ مادر

عزیزم

تقدیر و تشکر

حال که به لطف و یاری خداوند این پایان نامه به اتمام رسیده است بر خود لازم می دانم که از زحمات بی دریغ آقایان دکتر علی مرادزاده و دکتر محمد مهدی فاتح به عنوان اساتید راهنما که با رهنمودهایشان مرا در انجام هر چه بهتر این پایان نامه یاری فرمودند، صمیمانه تشکر نمایم. همچنین از کلیه اساتید ارجمند دانشکده خصوصاً آقای دکتر رضا خالو کاکائی، ریاست محترم دانشکده مهندسی معدن و ژئوفیزیک و آقای دکتر ایرج پیروز به دلیل کمک ها و راهنمایی های ارزشمندشان در طول تحصیل و در طی انجام این پایان نامه کمال تشکر و قدر دانی را دارم.

من در انجام این پایان نامه، خود را مدیون همه دوستانی می دانم که به هر نوعی از کمک و همراهی آن ها بهره برده ام، خصوصاً آقایان حسین شاه حسینی، علی رضا ترابی، مهندس حمید سرخیل و مهندس نادر زیاری که از هیچ گونه کمکی دریغ نکردند، ضمن تشکر از این دوستان از خداوند متان توفیق روزافزون برایشان خواستارم.

در پایان از کلیه دوستانی که مرا در انجام این پایان نامه یاری نموده اند، خصوصاً آقای امید سراج به دلیل کمک های ایشان در تهیه و طراحی تصاویر گرافیکی پایان نامه صمیمانه سپاسگزارم.

چکیده

روش مگنتوتلوریک یکی از روش‌های الکترومغناطیس با چشمه طبیعی می‌باشد که به منظور اکتشاف منابع معدنی، نفت و ژئوترمال استفاده می‌شود. اساس این روش، اندازه‌گیری تغییرات میدان‌های الکتریکی و مغناطیسی در سطح زمین می‌باشد که همانند بسیاری از روش‌های ژئوفیزیکی الکتریکی و الکترومغناطیسی توزیع مقاومت ویژه زمین را نسبت به عمق مشخص می‌سازد. مدل‌سازی وارون داده‌های مگنتوتلوریک و تخمین پارامترهای مدل، همانند سایر روش‌های ژئوفیزیکی مشکل‌ترین بخش تفسیر و پردازش این داده‌ها می‌باشد. از آن جایی که روش‌های معمول مدل‌سازی وارون غالباً نیاز به یک حدس اولیه در مورد پارامترهای مدل دارد لذا این امر مستلزم شناسایی برخی از پارامترهای مدل با استفاده از سایر روش‌های ژئوفیزیکی و یا اطلاعات زمین‌شناسی می‌باشد. از سوی دیگر روش‌های معمول مدل‌سازی وارون داده‌های ژئوفیزیکی به علت استفاده از الگوریتم‌های تکرار شونده بسیار وقت گیر می‌باشد. در مورد داده‌های مگنتوتلوریک، استفاده از این روش‌ها به منظور مدل‌سازی وارون، به علت خاصیت پخشی امواج الکترومغناطیسی منجر به حصول پاسخ کاملاً تقریبی می‌شود. با توجه به توانمندی‌های شبکه‌های عصبی به نظر می‌رسد که استفاده از آن راه حل مناسبی برای مدل‌سازی وارون داده‌های ژئوفیزیکی باشد. از این رو هدف از انجام این مطالعه بررسی توانمندی شبکه‌های عصبی در مدل‌سازی وارون دو بعدی داده‌های مگنتوتلوریک می‌باشد. به منظور نیل به این هدف ابتدا با توجه به شبه مقاطع رسم شده مقاومت ویژه و فاز داده‌های صحرائی، کلاس مدل انتخاب شد. پس از انجام مطالعات مفصل، یک شبکه تخمین‌گر به نام پرسپترون چند لایه با قانون فراگیری پس انتشار خطا طراحی و سپس از آن جهت مدل‌سازی وارون داده‌های مگنتوتلوریک استفاده شد.

برای آموزش و طراحی مناسب شبکه، در کل ۴۹۸ مدل آموزشی در کلاس مربوطه طراحی و پاسخ آن با استفاده از نرم‌افزار EMIXMT2D در ۱۱ فرکانس و ۹ ایستگاه، تولید شد. از میان مدل‌های فوق، ۳۹۸ مدل به منظور آموزش شبکه، ۱۰۰ مدل برای اعتبار سنجی و تعمیم شبکه و ۵۰ مدل برای آزمایش شبکه به طور تصادفی انتخاب شدند. همچنین به منظور بررسی میزان حساسیت شبکه به نوع داده‌های ورودی به شبکه، ۵ شبکه جداگانه با ورودی‌های مختلف آموزش داده شد و مشاهده شد که میزان دقت هر یک با توجه به فضای مدل‌های آموزشی متفاوت می‌باشد. بررسی نتایج

به دست آمده در تخمین پارامترهای مدل‌های آزمایشی در دو وضعیت بدون نویز و حضور چند درصد نویز، نشان می‌دهد که شبکه‌های آموزش داده شده از دقت قابل قبولی جهت مدل‌سازی داده‌های مگنتوتلوریک برخوردار می‌باشند. همچنین مقایسه پارامترهای مدل به دست آمده توسط یکی از شبکه‌های آموزش داده شده (شبکه MT) برای داده‌های صحرائی با نتایج حاصل از یکی از روش‌های معمول مدل‌سازی وارون (RRI) نشان می‌دهد که مدل‌ها و پارامترهای به دست آمده توسط دو روش فوق از شباهت و همخوانی خوبی برخوردار می‌باشند که این مطلب مؤید توانمندی شبکه طراحی شده برای مدل‌سازی معکوس داده‌های مورد نظر می‌باشد.

فهرست مطالب

عنوان	صفحه
تقدیر و تشکر	الف.....
چکیده	ب.....
فهرست مطالب	د.....
فهرست اشکال	ح.....
فهرست جداول	ک.....

فصل اول: مقدمه

۱-۱- ژئوفیزیک و نقش آن در اکتشاف	۲
۲-۱- روش‌های الکتریکی و الکترومغناطیسی	۳
۱-۲-۱- روش الکترومغناطیس	۴
۳-۱- مدل‌سازی داده‌های مگنتوتلوریک	۷
۴-۱- شبکه‌های عصبی مصنوعی و کاربرد آن در حل مسائل ژئوفیزیکی	۱۱
۱- ۵- هدف از مطالعه حاضر	۱۳
۱-۶- ساختار پایان‌نامه	۱۴

فصل دوم: تئوری روش مگنتوتلوریک

۱-۲- مقدمه	۱۷
۲-۲- اصول ریاضی و فیزیک حاکم بر امواج مگنتوتلوریک	۱۸
۱-۲-۲- تئوری الکترومغناطیس	۱۸
۲-۲-۲- عمق پوسته	۲۱
۳-۲- محاسبه مقاومت ویژه برای یک نیم فضای همگن	۲۱
۴-۲- محاسبه مقاومت ویژه ساختارهای یک بعدی	۲۳
۵-۲- ساختارهای دوبعدی	۲۴

۶-۲- معادلات موج الکترومغناطیسی برای هر قطبش	۲۵
۶-۲-۱- تانسور امپدانس	۲۶

فصل سوم: مدل سازی داده‌های ژئوفیزیکی

۳-۱- مقدمه	۳۰
۳-۲- هدف از انجام مدل سازی	۳۰
۳-۳- مدل سازی مستقیم	۳۲
۳-۴- مدل سازی معکوس	۳۳
۳-۴-۱- مدل سازی هموار	۳۴
۳-۴-۲- مدل سازی پارامتری	۳۸
۳-۵- معرفی نرم افزار EMIXMT2D	۴۱

فصل چهارم: مبانی شبکه‌های عصبی مصنوعی

۴-۱- مقدمه	۴۵
۴-۲- شبکه عصبی بیولوژیکی	۴۵
۴-۳- شبکه عصبی مصنوعی	۴۷
۴-۳-۱- وزن‌های ارتباطی	۴۸
۴-۴- تاریخچه شبکه‌های عصبی مصنوعی	۴۹
۴-۵- کاربرد شبکه‌های عصبی	۵۰
۴-۶- ساختار شبکه عصبی	۵۱
۴-۷- شبکه‌های پیشخور و پسخور	۵۴
۴-۷-۱- شبکه‌های پیشخور	۵۴
۴-۷-۲- شبکه‌های پسخور یا برگشتی	۵۴
۴-۸- دسته‌بندی شبکه‌ها بر اساس روش فراگیری	۵۵
۴-۸-۱- شبکه‌های دارای فراگیری با ناظر	۵۵

۴-۸-۲- شبکه‌های دارای فراگیری بدون ناظر.....	۵۵
۴-۹- کاربرد شبکه های عصبی در ژئوفیزیک.....	۵۶
۴-۱۰- پرسپترون با الگوریتم آموزشی پس انتشار خطا.....	۵۷
۴-۱۱- پارامترهای شبکه.....	۶۱
۴-۱۱-۱- تعداد لایه‌ها.....	۶۱
۴-۱۱-۲- تعداد نرون‌های پنهان.....	۶۲
۴-۱۱-۳- انتخاب نوع تابع محرک.....	۶۲
۴-۱۱-۴- مقدار دهی اولیه وزن‌ها و بایاس.....	۶۳
۴-۱۱-۴- نرخ یادگیری و گشتاور.....	۶۴
۴-۱۲- پیش پردازش و پس پردازش داده‌ها.....	۶۵
۴-۱۴- تعمیم شبکه.....	۶۶
۴-۱۴-۱- منظم سازی.....	۶۷
۴-۱۴-۲- توقف سریع.....	۶۷

فصل پنجم: وارونه‌سازی داده‌های مگنتوتلوریک با استفاده از شبکه‌های عصبی مصنوعی

۵-۱- مقدمه.....	۷۱
۵-۲- انتخاب شبکه مناسب.....	۷۲
۵-۳- انتخاب کلاس مدل.....	۷۳
۵-۴- تعیین پارامترهای ورودی و خروجی شبکه.....	۷۵
۵-۵- تولید نمونه های آموزشی.....	۷۶
۵-۶- طراحی شبکه عصبی.....	۷۹
۵-۷- روش کار.....	۸۲
۵-۷-۱- انتخاب روش نرمالایز کردن.....	۸۴
۵-۷-۲- انتخاب نوع تابع فعالیت برای نرون‌های لایه پنهان و لایه خروجی.....	۸۵
۵-۷-۳- انتخاب روش آموزشی مناسب.....	۸۶

۸۷	۵-۷-۲-الف - روش آموزشی لونیبرگ - مارکوآرت
۸۸	۵-۷-۴-تعمیم شبکه
۸۸	۵-۷-۵- تعیین تعداد لایه های پنهان و تعداد نرون های لایه پنهان
۹۱	۵-۷-۶- مقدار دهی اولیه برای وزن ها و بایاس ها
۹۷	۵-۸- اثر انتخاب نوع داده های ورودی
۹۷	۵-۸-۱- شبکه RES
۹۹	۵-۸-۲- شبکه PHASE
۱۰۱	۵-۸-۳- شبکه TE
۱۰۳	۵-۸-۴- شبکه TM
۱۰۶	۵-۸-۵- مقایسه و بررسی نتایج
۱۰۸	۵-۹- بررسی رفتار شبکه ها در حضور نویز
۱۲۱	۵-۱۰- آموزش شبکه برای تخمین یک پارامتر
۱۲۱	۵-۱۱- آزمایش شبکه با داده های صحرائی

فصل ششم: جمع بندی، نتیجه گیری و پیشنهادات

۱۲۵	۶-۱- جمع بندی
۱۲۶	۶-۲- نتیجه گیری
۱۲۸	۶-۳- پیشنهادات
۱۳۱	فهرست منابع
۱۴۱	پیوست (الف)

فهرست اشکال

عنوان	صفحه
شکل ۱-۱- سیستم اندازه‌گیری مؤلفه‌های میدان‌های مغناطیسی و الکتریکی	۵
شکل ۱-۲- مدل دو بعدی و دو وضعیت قطبش TE و TM	۲۵
شکل ۱-۳- مدل دریا و کف با مش بندی اجزای محدود	۳۳
شکل ۲-۳- نمودار اثر ریچ	۴۰
شکل ۳-۳- الگوریتم وارونه‌سازی نرم‌افزار EMIXMT2D	۴۲
شکل ۱-۴- نرون بیولوژیکی	۴۶
شکل ۲-۴- نرون مصنوعی	۴۸
شکل ۳-۴- الف: مدل یک نرون با چند ورودی ب: فرم خلاصه شده نرون با چند ورودی	۵۲
شکل ۴-۴- شبکه چند لایه‌ای	۵۳
شکل ۵-۴- شبکه پشخور یا برگشتی	۵۴
شکل ۶-۴- نگاشت از فضای p - بعدی ورودی به فضای m بعدی خروجی	۵۶
شکل ۷-۴- ساختار پرسپترون چند لایه پیشخور	۵۸
شکل ۸-۴- الف: تابع سیگموئیدی. ب: تابع خطی	۶۳
شکل ۹-۴- کمینه‌های محلی و سراسری	۶۴
شکل ۱۰-۴- تعمیم شبکه با استفاده از سری اعتبار سنجی	۶۸
شکل ۱-۵- شبه مقاطع مقاومت ویژه(بالا) و فاز (پایین) برای داده‌های صحرائی مد TE، مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد.	۷۴
شکل ۲-۵- شبه مقاطع مقاومت ویژه(بالا) و فاز (پایین) برای داده‌های صحرائی مد TM، مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد.	۷۴
شکل ۳-۵- کلاس مدل انتخاب شده(دو آنومالی هادی مدفون در یک نیم فضای همگن) مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد.	۷۵
شکل ۴-۵- پوشش محدوده تعریف شده برای پارامتر ششم (α)	۷۷
شکل ۵-۵- دو توده آنومالی مدفون در یک نیم فضای همگن همراه با فرکانس‌ها (پریودها) و ایستگاه‌های	

اندازه‌گیری	۷۸
شکل ۵-۶- فلوجارت تعیین تعداد نرون‌های پنهان	۸۹
شکل ۵-۷- نمودار تعیین تعداد نرون‌های پنهان شبکه MT	۹۰
شکل ۵-۸- ساختار شبکه MT به صورت خلاصه شده	۹۱
شکل ۵-۹- فلوجارت تعیین نقطه شروع مناسب	۹۳
شکل ۵-۱۰- نمودار تعیین نقطه شروع مناسب برای وزن‌ها و بایاس‌ها در شبکه MT	۹۴
شکل ۵-۱۱- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه MT به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ	۹۵
شکل ۵-۱۲- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه MT (خط چین)	۹۶
شکل ۵-۱۳- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه RES به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ	۹۸
شکل ۵-۱۴- مقایسه مدل واقعی (خط پر) و مدل تخمینی توسط شبکه RES (خط چین)	۹۹
شکل ۵-۱۵- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه PHASE به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ	۱۰۰
شکل ۵-۱۶- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه PHASE (خط چین)	۱۰۱
شکل ۵-۱۷- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه TE به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ	۱۰۲
شکل ۵-۱۸- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TE (خط چین)	۱۰۳
شکل ۵-۱۹- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه TM به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ	۱۰۴
شکل ۵-۲۰- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TM (خط چین)	۱۰۵
شکل ۵-۲۱- خطای RMS به ازای هر شبکه برای سری آزمایشی به همراه مقادیر عددی آن‌ها	۱۰۶
شکل ۵-۲۲- مقادیر خطای میانگین سری آزمایشی به ازای هر پارامتر برای شبکه‌های مختلف	۱۰۷
شکل ۵-۲۳- مقادیر خطای RMS به دست آمده برای شبکه‌های مختلف در حضور نویز	۱۰۹
شکل ۵-۲۴- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه MT (خط چین) در حضور ۵ درصد نویز	

- ۱۱۰.....
- شکل ۵-۲۵- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه RES (خط چین) در حضور ۵ درصد نویز
- ۱۱۱.....
- شکل ۵-۲۶- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه PHASE (خط چین) در حضور ۵ درصد نویز
- ۱۱۲.....
- شکل ۵-۲۷- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TE (خط چین) در حضور ۵ درصد نویز
- شکل ۵-۲۸- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TM (خط چین) در حضور ۵ درصد نویز
- ۱۱۴.....
- شکل ۵-۲۹- تغییرات مقادیر مقاومت ویژه و فاز در دو مد TE و TM، مربوط به یکی از نمونه‌های آزمایشی در غیاب نویز
- ۱۱۶.....
- شکل ۵-۳۰- تغییرات مقادیر مقاومت ویژه و فاز در دو مد TE و TM، مربوط به یکی از نمونه‌های آزمایشی در حضور ۵ درصد نویز
- ۱۱۶.....
- شکل ۵-۳۱- شبه مقطع خطای نسبی مقاومت ویژه مد TE (a)، شبه مقطع خطای نسبی فاز مد TE (b)، شبه مقطع خطای نسبی مقاومت ویژه مد TM (c)، شبه مقطع خطای نسبی فاز مد TM (d)، مدل تخمین زده شده توسط شبکه MT (e) در غیاب نویز
- ۱۱۹.....
- شکل ۵-۳۲- شبه مقطع خطای نسبی مقاومت ویژه مد TE (a)، شبه مقطع خطای نسبی فاز مد TE (b)، شبه مقطع خطای نسبی مقاومت ویژه مد TM (c)، شبه مقطع خطای نسبی فاز مد TM (d)، مدل تخمین زده شده توسط شبکه MT (e) در حضور ۵ درصد نویز
- ۱۲۰.....
- شکل ۵-۳۳- مدل تخمین زده شده برای داده‌های صحرائی با شبکه MT
- ۱۲۲.....
- شکل ۵-۳۴- مقطع مقاومت ویژه حاصل از وارونه سازی هموار ترکیبی دو مد TE و TM
- ۱۲۲.....

فهرست جداول

عنوان	صفحه
جدول ۵-۱- تغییرات خطای RMS با افزایش تعداد نمونه‌ها.....	۸۶
جدول ۵-۲- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه MT برای یکی از نمونه‌های آزمایشی.....	۹۶
جدول ۵-۳- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES برای یکی از نمونه‌های آزمایشی.....	۹۹
جدول ۵-۴- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه PHASE برای یکی از نمونه‌های آزمایشی.....	۱۰۱
جدول ۵-۵- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TE برای یکی از نمونه‌های آزمایشی.....	۱۰۳
جدول ۵-۶- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TM برای یکی از نمونه‌های آزمایشی.....	۱۰۵
جدول ۵-۷- درصد خطای میانگین سری آزمایشی به ازای شبکه‌های مختلف.....	۱۰۷
جدول ۵-۸- مقادیر RMS برای شبکه‌های مختلف در حضور نویز.....	۱۰۹
جدول ۵-۹- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه MT برای یکی از نمونه‌های آزمایشی.....	۱۱۰
جدول ۵-۱۰- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES برای یکی از نمونه‌های آزمایشی.....	۱۱۱
جدول ۵-۱۱- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه PHASE برای یکی از نمونه‌های آزمایشی.....	۱۱۲
جدول ۵-۱۲- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TE برای یکی از نمونه‌های آزمایشی.....	۱۱۳
جدول ۵-۱۳- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TM برای یکی از نمونه‌های آزمایشی.....	۱۱۴

- جدول ۵-۱۴- مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT. ۱۱۷.
- جدول ۵-۱۵- مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT با ۵ درصد نویز..... ۱۱۸.
- جدول ۵-۱۶- تخمین پارامترهای داده‌های صحرائی توسط شبکه MT..... ۱۲۱.

فصل اول:

مقدمه

۱-۱- ژئوفیزیک و نقش آن در اکتشاف

ژئوفیزیک دانش مطالعه زمین با استفاده از اندازه‌گیری خواص فیزیکی زمین می‌باشد، در یک برداشت ژئوفیزیکی هدف، تعیین محل ساختارهای زمین‌شناختی زیرزمینی و در صورت امکان اندازه‌گیری ابعاد و خواص فیزیکی مربوط به آن‌ها می‌باشد. در اکتشاف نفت، اطلاعات ساختاری مورد توجه است زیرا نفت با عوارض خاصی چون تاقدیس^۱ در سنگ‌های رسوبی ارتباط دارد. در کاوش‌های ژئوفیزیکی معدنی تأکید بر آشکارسازی و تعیین خواص فیزیکی می‌شود، هر چند کانسارهای معدنی^۲ نشانه‌های ژئوفیزیکی متمایز و قابل اندازه‌گیری از خود بروز می‌دهند ولی اغلب شکل نامنظم دارند و در سنگ‌هایی با ساختار پیچیده روی می‌دهند که تفسیر کمی دقیق داده‌ها را دشوار و یا غیر ممکن می‌سازند. در بررسی‌های ساختگاه^۳‌های مهندسی، ممکن است هم ساختار و هم خواص فیزیکی آن‌ها مورد توجه مهندسان باشد. در محل احداث ساختمان‌های بزرگ، اغلب تغییرات عمقی سنگ کف^۴ مورد نظر است و وقتی تحمل بارهای سنگین مد نظر باشد، ویژگی‌های مکانیکی روباره^۵ ممکن است اهمیت پیدا کند. یک برداشت ژئوفیزیکی شامل یک سری از اندازه‌گیری‌ها می‌باشد که معمولاً با یک طرح منظم روی سطح زمین، دریا، هوا و یا در یک چاه صورت می‌گیرد، این اندازه‌گیری‌ها ممکن است اندازه‌گیری تغییرات میدان‌های ایستا^۶ (گرادیان‌های پتانسیل الکتریکی، گرانشی یا مغناطیسی)

1-anticline
2-mineral deposits
3-site
4-bedrock
5-overburden
6-static

و یا خصوصیات میدان‌های موج به ویژه زمان سیر امواج کشسان^۱ (لرزه‌های^۲)، تغییرات دامنه و فاز امواج الکترومغناطیسی باشد، که تحت تأثیر خواص فیزیکی ساختار سنگ‌های زیر سطحی می‌باشند. از آن جایی که خواص فیزیکی تا حد زیادی توسط سنگ‌شناسی^۳ تعیین می‌گردد، ناپیوستگی‌ها در خواص فیزیکی، اغلب مربوط به مرزهای زمین شناختی می‌شود و لذا هرگونه مسأله ساختاری، به تفسیر این میدان‌ها در روی زمین بر حسب این ناپیوستگی‌ها برمی‌گردد. سهولت انجام این کار به عوامل بسیاری بستگی دارد که از آن میان، پیچیدگی ساختار و تباین^۴ خواص فیزیکی سنگ‌های سازنده آن ساختار، اهمیت ویژه‌ای دارند. تباین خواص فیزیکی سنگ‌های زیر زمین و همگنی^۵ آن‌ها در یک سازند خاص از عوامل مهمی است که باید در انتخاب روش ژئوفیزیکی در نظر گرفته شود. خواصی از سنگ‌ها که بیش از همه در اکتشافات ژئوفیزیکی از آن‌ها استفاده می‌شود عبارتند از کشسانی، مقاومت ویژه یا هدایت ویژه الکتریکی^۶، چگالی، نفوذپذیری مغناطیسی^۷، قطبش‌پذیری^۸ الکتریکی و خواص دیگری چون درجهٔ رادیواکتیویته که کمتر استفاده می‌شود [گیریفیث و کینگ^۹، ۱۹۸۸].

بر اساس نوع خاصیت فیزیکی اندازه‌گیری شده، روش‌های ژئوفیزیکی، به روش‌های لرزه‌ای، گرانی‌سنجی، مغناطیسی، الکتریکی و رادیومتری تقسیم می‌شوند که از میان آن‌ها روش‌های الکتریکی و الکترومغناطیسی در مطالعهٔ حاضر بیشتر مورد بررسی قرار می‌گیرد.

۲-۱- روش‌های الکتریکی و الکترومغناطیسی

اساس روش‌های الکتریکی در کاوش‌های ژئوفیزیکی، اندازه‌گیری خواص الکتریکی در سنگ‌ها می‌باشد. این روش‌ها بسیار متنوع و بر پایهٔ اندازه‌گیری اختلاف پتانسیل، جریان و تغییرات میدان‌های

1-elastic
 2-seismic
 3-lithology
 4-Contrast
 5-homogeneity
 6-electrical resistivity or conductivity
 7-sucseptibility
 8-polarizability
 9-Griffiths and King

الکتريکی و مغناطیسی در روش‌های الکترومغناطیسی می‌باشند، که به طور طبیعی و یا مصنوعی در زمین ایجاد می‌شوند.

تغییرات مقاومت ویژه یا هدایت ویژه سنگ‌ها و کانی‌ها از ویژگی‌هایی است که استفاده از این روش‌ها را امکان‌پذیر ساخته است. این روش‌ها بر اساس چشمه تولید جریان الکتريکی به دو گروه تقسیم می‌شوند.

روش‌های الکتريکی دارای چشمه طبیعی^۱ که از جریان‌های الکتريکی زمین بهره می‌گیرند که شامل روش‌های پتانسیل خودزا^۲، جریان تلوریک^۳ و مگنتوتلوریک^۴ (MT) می‌شوند و روش‌هایی که از چشمه مصنوعی^۵ بهره می‌گیرند که شامل: مقاومت ویژه الکتريکی، الکترومغناطیس و قطبش القایی^۶ می‌شوند.

۱-۲-۱- روش الکترومغناطیس

اساس روش الکترومغناطیس، القای جریان الکتريکی در ساختارهای زیر سطحی با استفاده از میدان‌های مغناطیسی متغیر می‌باشد که منشأ این میدان‌ها می‌تواند طبیعی و یا مصنوعی باشد. در روش‌های الکترومغناطیس با چشمه مصنوعی، جهت تولید میدان مغناطیسی از چشمه‌های مصنوعی مانند سیم بلند و یا سیم پیچ‌های بزرگ استفاده می‌شود. اما در روش‌های الکترومغناطیس چشمه طبیعی (MT) از تغییرات طبیعی در میدان مغناطیس زمین به عنوان چشمه میدان استفاده می‌شود. که در این تحقیق به بررسی روش مگنتوتلوریک پرداخته می‌شود.

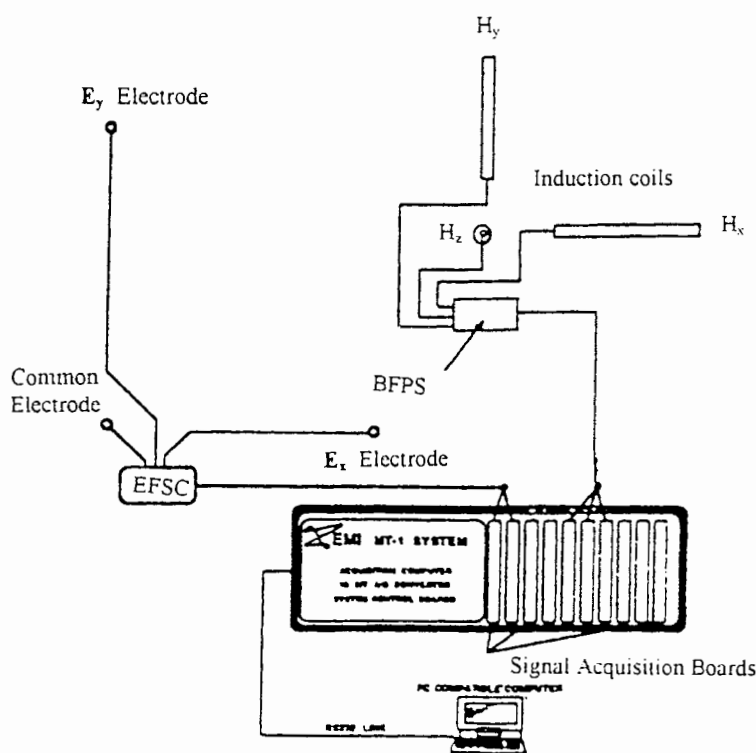
۱-۲-۲- روش مگنتوتلوریک

روش مگنتوتلوریک جهت به دست آوردن اطلاعات الکتريکی (توزیع فضایی مقاومت ویژه الکتريکی) از ساختارهای زیرزمینی استفاده می‌شود.

-
- 1-natural source
 - 2-Self Potential (SP)
 - 3-Telluric Current
 - 4-Magnetotelluric (MT)
 - 5-Artificial (controlled) source
 - 6-Induced Polarization

در این روش از امواجی که ناشی از هر گونه آشفتگی طبیعی در میدان مغناطیسی زمین می‌باشند، استفاده می‌شود. امواج MT گستره وسیعی از فرکانس‌ها را (از 10^{-5} هرتز تا آستانه شنوایی) در برمی‌گیرند. چشمه امواج MT دارای فرکانس‌های زیر ۱ هرتز، جریان‌های یونوسفری می‌باشد و امواج بالای ۱ هرتز ناشی از پدیده‌های الکتریکی در اتمسفر مانند طوفان‌های الکتریکی می‌باشند [مرادزاده، ۱۹۹۸]. این میدان‌های الکترومغناطیسی در سطح زمین همانند امواج تخت رفتار می‌کنند، بیشتر انرژی امواج تابشی در سطح زمین بازتاب و تنها بخش کوچکی از آن به طور عمود وارد زمین می‌شوند که باعث القای جریان‌های الکتریکی در داخل زمین و متعاقباً تولید میدان‌های مغناطیسی ثانویه می‌شوند.

این روش بر پایه اندازه‌گیری هم‌زمان تغییرات میدان‌های الکتریکی و مغناطیسی در سطح زمین می‌باشد. به منظور ثبت این اطلاعات، سه مؤلفه متعامد میدان مغناطیسی (H_z, H_y, H_x) در جهت‌های Z, Y, X و دو مؤلفه میدان الکتریکی (E_y, E_x) در جهت‌های y, x برای بازه‌های خاص فرکانس مطابق شکل ۱-۱ در یک نقطه از سطح زمین اندازه‌گیری می‌شود.



شکل ۱-۱- سیستم اندازه‌گیری مؤلفه‌های میدان‌های مغناطیسی و الکتریکی [مرادزاده، ۱۹۹۸].

داده‌ها در حوزه زمانی برداشت می‌شوند، اما با توجه به ساده بودن معادلات در حوزه فرکانس، داده‌ها با استفاده از تبدیل فوریه^۱ به حوزه فرکانسی منتقل می‌شوند. نسبت میدان‌های الکتریکی و مغناطیسی در هر فرکانس کمیت مختلطی به نام امپدانس الکتریکی ($Z_{ij} = \frac{E_i}{H_j}$) می‌باشد. با استفاده از امپدانس الکتریکی و رابطه زیر می‌توان مقاومت ویژه ظاهری و فاز امپدانس را در هر فرکانس به دست آورد.

$$\rho_{ij} = 0.2T |Z_{ij}|^2 \quad \Omega m \quad i, j = x, y$$

$$\varphi_{ij} = \tan^{-1} \left(\frac{\text{Im}(Z_{ij})}{\text{Re}(Z_{ij})} \right) \quad (1-1)$$

در رابطه فوق ρ_{ij} مقاومت ویژه ظاهری، T پریود اندازه‌گیری، φ_{ij} فاز امپدانس الکتریکی، $\text{Im}(Z_{ij})$ بخش موهومی امپدانس الکتریکی و $\text{Re}(Z_{ij})$ بخش حقیقی امپدانس الکتریکی می‌باشد. چگونگی محاسبه مقاومت ویژه با جزئیات بیشتر، در فصل دوم مورد بررسی قرار می‌گیرد.

عمق نفوذ امواج الکترومغناطیسی به فرکانس و مقاومت ویژه محیط بستگی دارد. با توجه به این که این امواج، بازه وسیعی از فرکانس‌ها را در برمی‌گیرند، دارای عمق نفوذ قابل توجهی (از چند ده متر تا چندین کیلومتر) می‌باشند. از این روش به دلیل داشتن عمق نفوذ خوب، حتی در مطالعه پوسته و گوشته فوقانی زمین نیز استفاده می‌شود. اما کاربرد عمده این روش در اکتشاف منابع معدنی، ژئوترمال و ذخایر هیدروکربوری به ویژه در مناطق کوهستانی و یا مناطقی که دارای روباره‌های ولکانیکی (بازالتی) است، می‌باشد. در این گونه موارد، روش‌های لرزه‌ای در اکتشاف ساختارهای زیرسطحی ناتوان است و بایستی از روش MT استفاده شود [مرادزاده، ۱۹۹۸]. با توجه به عدم نیاز به چشمه مصنوعی و حمل ابزار آلات سنگین، این روش نسبت به سایر روش‌های الکتریکی و الکترومغناطیسی برتری دارد. در ضمن این روش هیچ‌گونه آثار زیست محیطی در بر ندارد.

از نقطه ضعف‌های عمده این روش، طبیعت آشفته امواج در فرکانس‌های بالا و ضعیف بودن امواج چشمه در فرکانس‌های حدود ۱ Hz و ۲ KHz می‌باشد. راه حل این مشکل استفاده از روش CSAMT^۲ می‌باشد. در این روش از یک چشمه مصنوعی استفاده می‌شود که در نقطه‌ای دور از محل

1-Fourier transform

2-Controlled Source Audio Magnetotelluric

برداشت جهت تولید امواج در گستره فرکانس شنوایی استفاده می‌شود. نقطه ضعف دیگر روش MT مشکل جمع‌آوری داده در مناطق حاوی نویزهای الکتریکی می‌باشد، به منظور برطرف کردن این مشکل از روش $RRMT^1$ استفاده می‌شود که در این روش از ایستگاه‌های مبنا جهت اندازه‌گیری تغییرات میدان‌ها استفاده می‌شود [مرادزاده، ۱۹۹۸].

از مقادیر مقاومت ویژه و فاز امپدانس به دست آمده در تفسیر و مدل‌سازی ساختارهای زیر سطحی استفاده می‌شود.

۳-۱- مدل‌سازی داده‌های مگنتوتلوریک

هدف از مدل‌سازی داده‌ها، بازسازی یک مدل زمینی براساس داده‌های صحرائی (مشاهده‌ای) می‌باشد. مدل‌سازی‌ها به دو صورت مستقیم یا پیشرو^۲ و مدل‌سازی وارون یا معکوس^۳ صورت می‌گیرد. در مدل‌سازی مستقیم، مشخصات یک مدل زیر سطحی معلوم است که با استفاده از معادلات حاکم بر مسأله پاسخ مدل محاسبه می‌شود. مدل‌سازی وارون به منظور تعیین خصوصیات مدل زیر سطحی با استفاده از داده‌های مشاهده‌ای به کار می‌رود. مدل‌سازی مستقیم نسبت به مدل‌سازی وارون بسیار کند است ولی ریاضیات حاکم بر مدل‌سازی وارون بسیار پیچیده‌تر از مدل‌سازی مستقیم می‌باشد. با توجه به توانمندی مدل‌سازی وارون معمولاً از این مدل‌سازی برای یافتن مدلی که پاسخ آن بیشترین تطابق را با داده‌های مشاهده‌ای دارد، استفاده می‌شود [مرادزاده، ۱۹۹۸].

مدل‌سازی داده‌های MT می‌تواند به صورت یک، دو و سه بعدی^۴ صورت گیرد. یک ساختار ژئوالکتریکی یک بعدی، ساختاری است که در آن مقاومت ویژه تنها نسبت به عمق تغییر می‌کند، در ساختارهای ژئوالکتریکی دو بعدی، مقاومت ویژه نه تنها نسبت به عمق تغییر می‌کند بلکه در یک راستای افقی نیز تغییر می‌کند و در ساختارهای ژئوالکتریکی سه بعدی مقاومت ویژه علاوه بر تغییر نسبت به عمق، در دو راستای افقی نیز تغییر می‌کند.

1-Remote Reference Magnetotelluric

2-Forward modeling

3-Inverse modeling

4-one (1D), two (2D) and three dimensional (3D)

از آن جایی که داده‌های مشاهده‌ای اغلب آلوده به نویز هستند و با توجه به این که ما با داده‌های محدود مواجه هستیم و در فرآیند مدل‌سازی نیز مقداری خطا وجود دارد، از این رو پاسخ مدل‌سازی وارون یکتا^۱ نمی‌باشد. به بیان دیگر ممکن است پاسخ تعداد زیادی از مدل‌ها با یک سطح خطای معین، برازش خوبی با داده‌های مشاهده‌ای داشته باشند. در مجموع هدف از وارونه‌سازی داده‌ها به دست آوردن این مدل‌ها و انتخاب مدل‌های مناسب از بین آن‌ها می‌باشد که مفهوم فیزیکی داشته باشند. مدل‌سازی‌های وارون می‌توانند بر اساس نوع معیار مورد استفاده، جهت محدود کردن تعداد پاسخ‌های وارونه‌سازی که در برازش خوبی با داده‌ها قرار دارند و انتخاب مدل‌های مناسب، تقسیم بندی شوند. وارونه‌سازی پارامتری [چاپ و ووزوف^۲، ۱۹۷۷ و اوریستاگ‌لیو و وورسینگتون^۳، ۱۹۸۰] از استراتژی یافتن مدلی نزدیک به فرم ایده‌آل پیروی می‌کند. تفسیر در این روش شامل طراحی مدلی بر اساس اطلاعات اولیه زمین‌شناسی و محاسبه مقادیر هدایت ویژه به واسطه وارونه‌سازی حداقل مربعات می‌باشد. نتایج این روش‌ها به شیوه پارامتری کردن مدل، دقت و میزان دسترسی به اطلاعات اولیه زمین‌شناسی بستگی دارد.

روش دیگر، در نظر گرفتن تعداد زیادی پارامتر برای مسأله و مقید کردن وارونه‌سازی با استفاده از روش منظم‌سازی^۴ می‌باشد [تیخونوف^۵، ۱۹۶۳ a,b]. منظم‌سازی، کمینه کردن میزان خطای داده‌ها به طور همزمان می‌باشد. یک روش استاندارد، پیدا کردن هموارترین مدلی است که با داده‌ها برازش یابد. روش‌های مختلفی برای یافتن هموارترین مدل وجود دارد. [رودی^۶، ۱۹۸۹؛ دیگروت-هلدین^۷ و کانستابل^۸، ۱۹۹۰؛ اسمیت و بوکر^۹، ۱۹۹۱؛ یوچیدا^{۱۰}، ۱۹۹۳؛ رودی و مکی^{۱۱}، ۲۰۰۱].

بحث مدل‌سازی‌های پارامتری و هموار در فصل سوم به طور مفصل با استفاده از روابط ریاضی ارائه می‌شود.

-
- 1-Unique
 - 2-Jupp & Vozoff
 - 3-Oristaglio & Worthington
 - 4-Regularization
 - 5-Tikhonov
 - 6-Rodi
 - 7-de Groot-Heldin
 - 8-Constable
 - 9-Smith & Booker
 - 10-Uchida
 - 11-Mackie

در مقایسه دو روش وارونه‌سازی هموار و پارامتری می‌توان گفت که اعمال قید هموار، موجب پایداری الگوریتم وارونه‌سازی می‌شود هرچند که مدل‌های ناشی از این روش‌ها ممکن است، از نظر زمین‌شناسی دور از واقعیت باشند، این روش برای ساختارهای زیرسطحی شامل سازندهای زمین‌شناسی با مقاومت ویژه یکنواخت که توسط مرزهای مشخص از هم جدا شده‌اند دارای نتایج دور از دقت می‌باشد [دیگروت-هلدین و کانستیل، ۲۰۰۴].

روش مگنتوتلوریک مشابه سایر روش‌های الکترومغناطیسی شامل اندازه‌گیری میدان‌های الکترومغناطیسی با گستره فرکانسی پایین می‌باشند، این میدان‌ها دارای خاصیت پخشی^۱ هستند، به طوری که شرایط مرزی همانند پیوسته بودن چگالی جریان و میدان الکتریکی عملاً برقرار نیست، از این رو حل معادلات حاکم بر مسائل با استفاده از این شرایط مرزی، تقریبی می‌باشد به این ترتیب تعریف دقیق مرزها میسر نمی‌باشد، البته برای حل این مشکل روش‌های مختلفی ارائه شده است [مهانی و ژادانوف^۲، ۲۰۰۲؛ مارسللو-پاسکوال^۳ و همکاران، ۱۹۹۲؛ اسمیت و همکاران، ۱۹۹۹]. از جمله روش‌های وارونه‌سازی هموار می‌توان از روش‌های اوکام^۴ [کانستیل و همکاران، ۱۹۸۷] و RRI^۵ [اسمیت و بوکر، ۱۹۹۱، ۱۹۸۸] نیز نام برد. در روش اوکام پارامتر ناهمواری^۶ به عنوان قید هموارسازی به مسأله اعمال می‌شود همان‌طور که در فصل سوم اشاره می‌شود، این قید به صورت مربع مجموع مشتق اول یا دوم پارامترهای مدل نسبت به عمق تعریف می‌شود. در این روش هدف کمینه کردن تابع زیر می‌باشد.

$$U = R(m) + \mu^{-1}(\chi^2 - \chi^{*2}) \quad (2-1)$$

ترم اول در سمت راست رابطه فوق قید ناهمواری و ترم بعدی خطا می‌باشد که به وسیله ضریب لاگرانژ^۶ (μ^{-1}) وزن‌دهی شده است، که در آن χ^2 اختلاف بین پاسخ مدل اولیه و داده‌های مشاهده‌ای می‌باشد و χ^{*2} نیز سطح قابل قبول خطا می‌باشد [کانستیل و همکاران، ۱۹۸۷].

1-Diffusive
 2-Mehanee & Zhdanov
 3-Marcuello-Pascuel
 4-Occam
 5-Rapid Relaxation Inversion
 6-Roughness

در ساختارهای دوبعدی پارامتر ناهموازی در دو بعد تعریف می‌شود [مرادزاده، ۱۹۹۸]. اما این روش در وارونه‌سازی مدل‌هایی که دارای تباین مقاومت ویژه زیادی می‌باشند، دچار مشکل می‌شود که البته اخیراً روشی برای حل این معضل نیز ارائه شده است [کانستیبیل و دیگرگوت-هدلین، ۲۰۰۴]. روش RRI نیز یکی از روش‌های وارونه‌سازی هموار می‌باشد که این روش بر اساس کمینه کردن مشتق دوم رسانندگی در دو جهت قائم و افقی می‌باشد [مرادزاده ۱۹۹۸]. در وارونه‌سازی هموار، تعیین دقیق محل آنومالی‌ها و مرز جدایش آنها مشکل و بعضاً امکان‌پذیر نیست. از این رو از مدل‌سازی پارامتری استفاده می‌شود [مرادزاده و عرب امیری، ۱۳۸۳]. در فرآیند وارونه‌سازی داده‌ها، محاسبه پاسخ مدل اولیه توسط مدل‌سازی مستقیم صورت می‌گیرد در این مدل‌سازی با حل معادلات حاکم بر مسأله پاسخ مدل محاسبه می‌شود. هدف از حل معادلات امواج مگنتوتلوریک محاسبه میدان‌های الکتریکی و مغناطیسی می‌باشد با استفاده از این میدان‌ها می‌توان امپدانس الکتریکی و متعاقباً مقاومت ویژه و فاز را در هر فرکانس به دست آورد. روش به دست آوردن معادلات حاکم بر امواج مگنتوتلوریک در دو بعد به صورت زیر می‌باشد که در فصل دوم بررسی می‌شود.

$$\nabla^2 \begin{pmatrix} E \\ H \end{pmatrix} - j\omega\mu\sigma \begin{pmatrix} E \\ H \end{pmatrix} = 0 \quad (3-1)$$

در معادله فوق E و H مؤلفه‌های میدان‌های الکتریکی و مغناطیسی، σ هدایت ویژه الکتریکی، μ نفوذ پذیری مغناطیسی، ω فرکانس زاویه‌ای و $j = \sqrt{-1}$ می‌باشند. حل این معادلات تنها با استفاده از روش‌های عددی همانند اجزای محدود^۱ و تفاضل‌های محدود^۲ امکان‌پذیر است.

روش‌های فوق با استفاده از مش‌بندی مدل (تقسیم بندی مدل با استفاده از خطوط طولی و عرضی) و شرایط مرزی معادلات مذکور را حل می‌کنند. اما بدیهی است با افزایش بعد، مش‌بندی‌ها پیچیده‌تر و تعداد معادلات نیز افزایش می‌یابد، در نتیجه زمان محاسبات افزایش و دقت کاهش

1-finite element
2-finite difference

می‌یابد. در حل معادلات، روش اجزای محدود دارای تقریب خوبی می‌باشد و به دلیل مش‌بندی خوب و قابل انعطاف آن، مناسب برای مسائل با تغییرات شدید توپوگرافی می‌باشد اما ریاضیات سنگینی دارد، از این رو تعمیم آن به مسائل دو بعدی و سه بعدی علاوه بر پیچیدگی، موجب تشکیل دستگاه‌های بسیار بزرگی از معادلات خواهد شد، در نتیجه زمان زیادی را برای پردازش و مدل‌سازی صرف می‌کند. روش تفاضل‌های محدود، علی‌رغم صرف زمان کمتر دارای دقت کمتری نیز، مخصوصاً در مدل‌های دارای توپوگرافی، می‌باشد [گرامی زادگان، ۱۳۸۲]. روش‌های عددی فوق در حل معادلات مربوط به مدل‌های سه بعدی عملاً ناتوان می‌باشند، در مورد این مدل‌ها از روش معادلات انتگرالی^۱ استفاده می‌شود، ولی این روش از دقت بسیار پایینی برخوردار است و بیشتر برای مدل‌های بسیار ساده استفاده می‌شود [مکی و همکاران^۲، ۱۹۹۳].

با توجه به ماهیت پخشی امواج مگنتوتلوریک و محدودیت‌های روش‌های فوق، پاسخ معادلات در هر نقطه به طور تقریبی محاسبه می‌شود پس در نهایت این روش منجر به تولید یک مدل تقریبی می‌شود. با توجه به مشکلات روش‌های عددی در حل معادلات به ویژه در مدل‌های پیچیده و زمان‌بر بودن الگوریتم‌های معمول وارونه‌سازی داده‌ها، یافتن روش‌های جدیدی برای مدل‌سازی داده‌های MT همواره مورد توجه محققین می‌باشد. با توجه به غیر خطی بودن رابطه بین پارامترهای مدل و پاسخ مدل و توانمندی‌های شبکه‌های عصبی مصنوعی، به نظر می‌رسد که این شبکه‌ها بتوانند راه حلی برای مسأله وارونه‌سازی داده‌های مقاومت ویژه و فاز در روش MT ارائه کنند. با توجه به این که در این زمینه کار چندانی صورت نگرفته است انجام موارد فوق در این مطالعه برای ساختارهای دو بعدی مورد بررسی قرار می‌گیرد.

۴-۱- شبکه‌های عصبی مصنوعی و کاربرد آن در حل مسائل ژئوفیزیکی

شبکه‌های عصبی مصنوعی، الگوی بسیار ساده‌ای از سیستم عصبی انسان می‌باشند. آن‌ها دارای قابلیت یادگیری هستند. سیستم یادگیری آن‌ها مشابه سیستم یادگیری انسان می‌باشد. شبکه‌های عصبی مصنوعی با دریافت الگوهای ورودی قادر به استخراج مشخصات آن‌ها و نسبت دادن یک کلاس

1-Integral Equation

2-Mackie et al

خاص به آنها می‌باشند به این ترتیب در خلال فرآیند آموزش، آنها مشخصات الگوها و کلاس مورد نظر را با تطبیق دادن یک سری ضرایب وزنی به خاطر می‌سپرنند. چنانچه آموزش به درستی صورت گیرد آنها قادر به دسته‌بندی^۱ الگوها می‌باشند.

آنچه که معمولاً در زمینه ژئوفیزیک از شبکه می‌خواهیم دسته‌بندی داده‌ها و الگوها و یا محاسبات یک رگرسیون است، وقتی که بخواهیم یک مقدار خاصی را بر پایه یک الگوی ورودی تخمین بزنیم [پولتون^۲، ۲۰۰۱].

شبکه‌های عصبی برای حل مسائل غیر خطی بسیار مناسب می‌باشند، آنها تفسیر و پردازش همزمان داده‌ها را میسر می‌سازند. از این رو شبکه‌های عصبی کاربرد فراوانی در مسائل ژئوفیزیکی به ویژه در دهه اخیر یافتند [اسپیشاک و پوپوف^۳، ۲۰۰۰]. شبکه‌های عصبی به طور موفقیت‌آمیزی در شاخه‌های مختلف ژئوفیزیکی مانند: الکترومغناطیس [پولتون و همکاران، ۱۹۹۲]، مگنتوتلوریک [اسپیشاک و پوپوف، ۲۰۰۰]، وارونه‌سازی داده‌های لرزه [روت و تارانتولا^۴، ۱۹۹۴؛ لانگر و همکاران^۵، ۱۹۹۶؛ کلدرون - ماسیاس و همکاران^۶، ۱۹۹۸]، تحلیل نمودارهای چاه‌پیمایی [هوانگ و همکاران^۷، ۱۹۹۶]، دکانولوشن لرزه [وانگ و مندال^۸، ۱۹۹۲؛ کلدرون - ماسیاس و همکاران، ۱۹۹۷]، دسته‌بندی حادثه‌ها [دوولا و همکاران^۹، ۱۹۹۰؛ رومئو^{۱۰}، ۱۹۹۴] به کار گرفته شده است. در زمینه کاربرد شبکه‌های عصبی در ژئوفیزیک، کار زیادی در ایران انجام نشده است. از بین تعداد کمی کار صورت گرفته، می‌توان به کارهای انجام شده در زمینه مدل‌سازی دو بعدی معکوس داده‌های مقاومت ویژه [گرامی زادگان، ۱۳۸۲] و داده‌های چاه‌پیمایی [بخشی، ۱۳۸۲] اشاره کرد. بررسی‌ها نشان می‌دهد که متأسفانه استفاده از شبکه‌های عصبی در حل مسائل روش MT بسیار محدود و به دو مورد زیر خلاصه می‌شود.

-
- 1-Classification
 - 2-Poulton
 - 3- Spichak and popova
 - 4-Roth & Tarantola
 - 5-Langer et al
 - 6-Calderon-Macias et al
 - 7-Huang et al
 - 8-Wang & Mendal
 - 9-Dowla et al
 - 10-Romeo

مورد اول مربوط به مطالعه محدود جهت تولید و ساختن مدل اولیه جهت مدل سازی معکوس یک بعدی داده‌های MT در حوزه زمان می‌باشد که در این تحقیق از یک شبکه عصبی به نام RBF^1 استفاده شده است [هیدالگو و همکاران^۲، ۱۹۹۴].

مورد دوم تحقیق دیگری است که در زمینه مدل سازی داده‌های AMT^3 (مگنتوتلوریک در گستره فرکانسی شنوایی) با استفاده از شبکه‌های عصبی انجام شد [اسپیشاک و پوپوف، ۲۰۰۰]. در مطالعه فوق پاسخ یک مدل ساده سه بعدی مرکب از یک دایک شیب‌دار در زمین دولایه‌ای جهت آموزش شبکه مورد استفاده قرار گرفت و از شبکه آموزش داده شده برای تخمین پارامترهای شش‌گانه مدل استفاده گردید. علی‌رغم این که مدل انتخاب شده بسیار ساده بود ولی با توجه به محدودیت‌های اشاره شده در محاسبه پاسخ مدل‌های سه بعدی، نتایج به دست آمده طبیعتاً نمی‌توانستند خیلی دقیق باشند. با توجه به پیچیدگی‌های روابط حاکم بر معادلات روش MT و طبعاً پیچیدگی‌های موجود در روابط حاکم بر داده‌ها و نبود مطالعات جامع در زمینه به کارگیری روش شبکه‌های عصبی در مدل سازی داده‌های MT و مقایسه نتایج آن‌ها با دیگر روش‌های مدل سازی، انجام مطالعات در این زمینه از اهمیت زیادی برخوردار می‌باشد. از این رو در این پایان نامه سعی شده است تا توانمندی‌های شبکه عصبی در وارونه‌سازی دو بعدی داده‌های MT بیشتر مورد بررسی قرار گیرد.

۱-۵ - هدف از مطالعه حاضر

هدف از انجام این مطالعه بررسی توانمندی‌های شبکه عصبی در حل مشکلات وارونه‌سازی دو بعدی داده‌های مگنتوتلوریک می‌باشد. جهت نیل به این هدف، ابتدا کلاس مدل مورد نظر با توجه به شبه مقاطع داده‌های مشاهده‌ای مقاومت ویژه و فاز به صورت دو توده هادی مدفون در یک نیم فضای همگن انتخاب شد. هدف از وارونه‌سازی داده‌های مگنتوتلوریک در این کلاس، تخمین (۹ پارامتر مشتمل بر) پارامترهای هندسی، نسبت مقاومت ویژه و تعیین محل آنومالی‌ها می‌باشد. پس از انتخاب کلاس، نوبت به انتخاب یک شبکه مناسب می‌رسد که با توجه به کارهای صورت گرفته، شبکه

1-Radial Basis Function

2-Hidalgo et al

3-Audio Magnetotelluric

پرسپترون چند لایه^۱ (MLP) انتخاب شد. به منظور آموزش شبکه، مدل‌های آموزشی در ۱۱ فرکانس و ۹ ایستگاه با استفاده از یک نرم‌افزار مدل‌سازی دو بعدی داده‌های مگنتوتلوریک (EMIXMT2D) تولید شد [راهنمای نرم‌افزار EMIXMT2D، ۱۹۹۹]. این مدل‌ها با تغییر محل آنومالی‌ها نسبت به یکدیگر و ابعاد آن‌ها و تغییر همزمان مقاومت ویژه زمینه تولید شدند.

سپس با استفاده از مدل‌های آموزشی ساختار شبکه طراحی شد، که چگونگی انجام آن در فصل پنجم بررسی می‌شود. پس از طراحی، با داده‌های فاز و مقاومت ویژه که توسط نرم‌افزار EMIXMT2D برای تعداد زیادی از مدل‌های مصنوعی تولید شده‌اند، شبکه مورد آموزش لازم قرار گرفته است و سپس با انجام یک سری بررسی‌ها و آزمایشات، اعتبار صحت محاسبات توسط شبکه طراحی شده مورد بررسی قرار گرفت. پس از حصول اطمینان از صحت محاسبات، از شبکه طراحی شده جهت مدل‌سازی معکوس داده‌های واقعی MT استفاده شد و نتایج به دست آمده توسط شبکه با نتایج دیگر روش‌های مدل‌سازی معکوس مورد مقایسه قرار گرفته است. برنامه‌های مربوط به طراحی ساختار شبکه و آموزش آن با استفاده از نرم‌افزار matlab نوشته و تهیه گردید.

۶-۱ - ساختار پایان‌نامه

این پایان‌نامه شامل شش فصل می‌باشد. در این فصل مقدمه‌ای از ضرورت، هدف و چگونگی انجام این مطالعه ارائه شده است. در فصل دوم منشاء و خصوصیات میدان‌های مگنتوتلوریک، روابط ریاضی و فیزیکی حاکم بر آن‌ها و چگونگی محاسبه امپدانس و مقاومت ویژه الکتریکی در روش MT مورد بررسی قرار می‌گیرد. فصل سوم شامل بررسی روش‌های مختلف مدل‌سازی داده‌های ژئوفیزیکی و معرفی نرم‌افزار EMIXMT2D جهت تولید داده‌های آموزشی می‌باشد. در فصل چهارم به منظور آشنایی با شبکه‌های عصبی، اصول و مبانی شبکه‌های عصبی به همراه الگوریتم آموزشی آن‌ها بیان می‌شود. فصل پنجم شامل نحوه انتخاب نوع شبکه و تولید مدل‌های آموزشی، طراحی ساختار و آموزش شبکه، بررسی نتایج به دست آمده توسط شبکه برای ورودی‌های مختلف و بررسی رفتار شبکه در برابر اضافه کردن نویز به داده‌های ورودی و مقایسه نتایج تخمین پارامترهای مدل توسط شبکه

1-Multilayer Perceptron

برای داده‌های صحرایی و وارونه‌سازی آن‌ها با یکی از روش‌های معمول مدل‌سازی هموار می‌باشد و نهایتاً نتایج به دست آمده به همراه پیشنهادات در قالب فصل ششم ارائه می‌شود.

فصل دوم:

تئوری روش مگنتوتلوریک

تئوری روش مگنتوتلوریک

۱-۲- مقدمه

همان طور که قبلاً اشاره شد، روش مگنتوتلوریک یکی از روش‌های ژئوفیزیکی الکترومغناطیسی با چشمه طبیعی می‌باشد که از اوایل سال ۱۹۵۰ در اکتشافات نفت، منابع معدنی و منابع ژئوترمال به کار گرفته شد. اساس این روش، اندازه‌گیری تغییرات میدان‌های الکتریکی و مغناطیسی در سطح زمین می‌باشد که همانند بسیاری از روش‌های ژئوفیزیکی الکتریکی و الکترومغناطیسی توزیع مقاومت ویژه زمین را نسبت به عمق مشخص می‌کند [دوبرین و ساویت، ۱۹۸۸]. امواج مگنتوتلوریک طیف وسیعی از فرکانس‌ها را در بر می‌گیرد که از گستره فرکانسی 0.01 Hz تا 10000 Hz جهت کارهای اکتشافی استفاده می‌گردد، کاربرد اصلی این روش در اکتشاف نفت است به ویژه در مناطقی که روش لرزه‌ای انعکاسی به دلیل توپوگرافی شدید، خیلی گران می‌باشد و یا به دلیل حضور روباره‌های بازالتی و یا سایر ولکانیک‌ها، دچار مشکل می‌شود. گذشته از اکتشاف نفت، حساسیت روش MT به آنومالی‌های هادی، آن را جهت اکتشاف منابع ژئوترمال یک روش مناسب کرده است [ووزوف، ۱۹۹۱]. منشأ امواج MT ناشی از فعالیت‌های خورشیدی، آذرخش‌ها، شفق‌های قطبی، حرکت‌های نسبی زمین، خورشید و ماه و همچنین میدان‌های ناشی از گردبادهای بزرگ، می‌باشد [نقی‌زاده، ۱۳۸۲].

امواج ایجاد شده تقریباً بدون تضعیف، پوسته نارسانای هوا را طی کرده و به سطح زمین برخورد می‌کنند، بخش زیادی از میدان‌های تابشی در سطح زمین بازتاب می‌شود و تنها بخش کوچکی از آن به صورت امواج تخت به داخل زمین نفوذ می‌کنند. برای اندازه‌گیری در یک ایستگاه سه مؤلفه میدان

مغناطیسی (H_z, H_y, H_x) در راستای x, y, z و دو مؤلفه میدان الکتریکی (E_y, E_x) در راستای x, y به ترتیب توسط سه مگنتومتر (بیشتر سیم‌پیچ‌های مخصوص) و یک جفت الکتروود پتانسیل (مطابق شکل ۱-۱) ثبت می‌شوند و پس از عبور از یک سری تقویت‌کننده‌ها و فیلترها به همراه سیستم‌های دیجیتالی و پردازنده‌های مناسب در حافظه دستگاه ذخیره می‌شوند. داده‌ها در حوزه زمان برداشت می‌شوند ولی معمولاً پردازش در حوزه فرکانسی صورت می‌گیرد، چون تئوری آن ساده تر از حوزه زمانی است، پس پردازش داده‌ها با تبدیل فوریه آغاز می‌شود. در بسیاری از سیستم‌های امروزی، پردازش در همان زمان برداشت صورت می‌گیرد. نسبت میدان‌های ثبت شده، کمیتی از نوع امپدانس الکتریکی است. با تعیین امپدانس $Z_{ij} = \frac{E_i}{H_j}$ در یک سری از فرکانس‌ها می‌توان یک طیف فرکانسی برای امپدانس به دست آورد. در فرکانس‌های نسبتاً زیاد به علت اثر عمق پوسته، امپدانس تنها اطلاعاتی از لایه‌های سطحی می‌دهد و در فرکانس‌های پایین، این اطلاعات ناشی از لایه‌های عمقی می‌باشد، بنابراین مشاهده می‌شود که روش MT در یک ایستگاه به صورت یک سونداژ عمقی عمل می‌کند. در مرحله بعدی، امپدانس‌های محاسبه شده به صورت مقاومت ویژه بر حسب عمق تفسیر می‌شود. مرحله پایانی استفاده از روش‌های مدل‌سازی جهت تعیین ساختار زیر سطحی می‌باشد [نقی‌زاده، ۱۳۸۲].

۲-۲- اصول ریاضی و فیزیک حاکم بر امواج مگنتوتلوریک

بدیهی است که جهت توصیف انتشار امواج MT، تئوری الکترومغناطیس مورد استفاده قرار می‌گیرد. این امر مستلزم آگاهی از معادلات ماکسول می‌باشد، در واقع این معادلات می‌توانند آغازی برای دستیابی بر چگونگی کاربرد میدان‌های الکترومغناطیسی در اکتشافات ژئوفیزیکی برای تعیین خواص الکتریکی و مغناطیسی زمین باشد.

۲-۲-۱- تئوری الکترومغناطیس

قوانین مربوط به الکتریسیته و مغناطیس توسط ماکسول به صورت چهار معادله زیر مطرح شد.

$$\nabla \times E = -\frac{\partial B}{\partial t} \quad \text{قانون فارادی} \quad (1-2)$$

$$\nabla \times H = J + \frac{\partial D}{\partial t} \quad \text{قانون آمپر} \quad (2-2)$$

$$\nabla \cdot B = 0 \quad \text{عدم وجود تک قطب مغناطیسی} \quad (3-2)$$

$$\nabla \cdot D = \rho \quad \text{قانون کولمب} \quad (4-2)$$

که در این معادلات $B \left[\frac{wb}{m^2} \text{ Tesla} \right]$ القاء مغناطیسی، $H \left[\frac{A}{m} \right]$ شدت میدان مغناطیسی، $E \left[\frac{V}{m} \right]$

شدت میدان الکتریکی، $D \left[\frac{C}{m^2} \right]$ جابه‌جایی الکتریکی، $J \left[\frac{A}{m^2} \right]$ چگالی جریان الکتریکی، $\rho \left[\frac{C}{m^3} \right]$

چگالی حجمی بار الکتریکی و t زمان می‌باشد [مرادزاده، ۱۹۹۸].

در روش MT مؤلفه‌های افقی میدان الکتریکی یعنی مؤلفه‌های موازی سطح زمین اندازه‌گیری می‌شوند بنابراین هیچ بار سطحی در مرز هوا و زمین ظاهر نخواهد شد پس چگالی جریان ثابت می‌ماند. برای محاسبه معادلات حاکم بر انتشار امواج MT، از معادله پیوستگی (معادله ۲-۵) و معادلات مربوط به یک محیط خطی (معادلات ۲-۶، ۲-۷ و ۲-۸) استفاده می‌شود [ریتز و همکاران ۱۳۷۶].

$$\nabla \times J + \frac{\partial \rho}{\partial t} = 0 \quad \text{معادله پیوستگی} \quad (5-2)$$

$$J = \sigma E \quad \text{قانون اهم} \quad (6-2)$$

$$D = \varepsilon E \quad (7-2)$$

$$B = \mu H \quad (8-2)$$

که در معادلات فوق $\sigma \left[\frac{S}{m} \right]$ هدایت الکتریکی یا رسانندگی، $\varepsilon \left[\frac{F}{m} \right]$ گذردهی الکتریکی و $\mu \left[\frac{Tm}{A} \right]$

نفوذپذیری مغناطیسی می‌باشد. چنانچه تغییرات زمانی میدان‌ها را هارمونیک ($e^{j\omega t}$) در نظر بگیریم

معادلات ماکسول با فرض $\nabla \cdot D = 0$ و $\nabla \cdot H = 0$ به صورت زیر در می‌آیند.

$$\nabla^2 \begin{bmatrix} E \\ H \end{bmatrix} + (\mu\varepsilon\omega^2 - j\mu\sigma\omega) \begin{bmatrix} E \\ H \end{bmatrix} = 0 \quad (9-2)$$

و یا

$$\nabla^2 \begin{bmatrix} E \\ H \end{bmatrix} + K^2 \begin{bmatrix} E \\ H \end{bmatrix} = 0 \quad (10-2)$$

که در آن

$$K^2 = \mu\varepsilon\omega^2 - j\mu\sigma\omega \quad (11-2)$$

می‌باشد. معادلات (10-2) معادله هلمهولتز^۱ می‌باشند قسمت اول رابطه (11-2) مربوط به جریان‌های جابه‌جایی و قسمت دوم مربوط به جریان‌های رسانش می‌باشد. معادله هلمهولتز برای سنگ‌های با رسانندگی پایین و همچنین برای هوا به معادله لاپلاس تبدیل می‌شود.

$$\nabla^2 \begin{bmatrix} E \\ H \end{bmatrix} = 0 \quad (12-2)$$

همچنین برای یک رسانای خوب قسمت حقیقی معادله (9-2) حذف و تنها قسمت

موهومی آن باقی می‌ماند و معادله هلمهولتز تبدیل به معادله پخشی می‌شود.

$$\nabla^2 \begin{bmatrix} E \\ H \end{bmatrix} - j\mu\sigma\omega \begin{bmatrix} E \\ H \end{bmatrix} = 0 \quad (13-2)$$

در هر دو مورد فوق ملاحظه می‌شود که فاکتور مربوط به جریان‌های جابه‌جایی با توجه به گستره فرکانسی مورد استفاده در روش MT قابل صرف‌نظر کردن می‌باشد [تلفورد و همکاران^۲، ۱۹۹۰]. حل معادله (13-2) بسیار مشکل می‌باشد مگر آن که موج را قطبیده تخت^۳ در نظر بگیریم. تخت بودن امواج در گستره فرکانسی مورد نظر MT توسط محققین بسیاری بررسی و مورد تایید قرار گرفته است [دوبرین و ساویت، ۱۹۸۸]. چنانچه میدان الکترومغناطیسی تختی را در نظر بگیریم که در آن موج در جهت Z منتشر شود، صفحه xy صفحه قطبش باشد و مؤلفه‌های میدان‌ها را به صورت زیر فرض کنیم.

$$H = H_y(z, t)$$

1-Helmholtz
2-Telford et al
3-Plane wave

$$E = E_x(z, t)$$

پاسخ معادله (۱۳-۲) به این صورت است.

$$A = A_0 e^{-j(kz - \omega t)} = A_0 e^{-\alpha z} e^{j(\omega t - \alpha z)} \quad (14-2)$$

$$\alpha = \left(\frac{\omega \mu \sigma}{2} \right)^{1/2}, k = (1 - j)\alpha$$

A_0 بزرگی میدان‌های E, H در سطح زمین می‌باشد. از نظر فیزیکی تنها بخش حقیقی جواب معادله مورد نظر ماست بنابراین

$$A = A_0 e^{-\alpha z} \cos(\omega t - \alpha z) \quad (15-2)$$

ملاحظه می‌شود که دامنه موج به طور نمایی نسبت به عمق کاهش می‌یابد [مرادزاده، ۱۹۹۸].

۲-۲-۲- عمق پوسته^۱

به عنوان معیاری برای نفوذ امواج الکترومغناطیسی در داخل زمین فاکتوری به نام عمق پوسته تعریف می‌شود. عمق پوسته در یک زمین همگن و همسانگرد عمقی است که در آن دامنه موج به e^{-1} مقدارش در سطح زمین کاهش یابد.

$$\delta = 1/\alpha$$

$$\delta = \left(\frac{2}{\omega \mu \sigma} \right)^{1/2} \approx 503 \sqrt{\frac{\rho}{f}} \quad (16-2)$$

در معادله فوق ρ [Ωm] مقاومت ویژه زمین، f [HZ] فرکانس موج و δ [m] عمق پوسته می‌باشد.

۲-۳- محاسبه مقاومت ویژه برای یک نیم فضای همگن

موج الکترومغناطیسی را در نظر بگیرید که در راستای Z منتشر و در صفحه xy قطبیده شده باشد و همچنین تغییرات زمانی آن به صورت $e^{j\omega t}$ باشد.

$$E = (E_x, E_y, 0)$$

$$Z(\omega) = \frac{E_y}{H_x} = -\frac{E_x}{H_y} \quad (25-2)$$

با توجه به این که امپدانس یک کمیت مختلط می‌باشد، بنابراین علاوه بر اندازه، دارای فاز نیز می‌باشد. φ به عنوان فاز امپدانس Z ، همان اختلاف فاز میدان‌های E و H می‌باشد که در مورد یک نیم فضای همگن این اختلاف فاز $\frac{\pi}{4}$ است [تلفورد و همکاران، ۱۹۹۰].

۴-۲- محاسبه مقاومت ویژه ساختارهای یک بعدی

ساختارهای یک بعدی شامل چند لایه همگن افقی می‌باشد که مقاومت ویژه در آن‌ها تنها با عمق تغییر می‌کند. برای چنین ساختاری تغییرات افقی میدان‌ها صفر است. این مورد کاملاً شبیه مدل نیم فضای همگن می‌باشد. در ساختارهای یک بعدی نیز امپدانس مستقل از جهت اندازه‌گیری می‌باشد، بنابراین برای چنین ساختاری، امپدانس، فاز و مقاومت ویژه به صورت زیر می‌باشد.

$$Z(\omega) = \frac{E_x(\omega)}{H_y(\omega)} = -\frac{E_y(\omega)}{H_x(\omega)} \quad (26-2)$$

$$\rho_{xy} = 0.2T |Z_{xy}|^2, \quad \varphi_{xy} = \tan^{-1} \left[\frac{\text{Im}(Z_{xy})}{\text{Re}(Z_{xy})} \right] \quad (27-2)$$

$$\rho_{yx} = 0.2T |Z_{yx}|^2, \quad \varphi_{yx} = \tan^{-1} \left[\frac{\text{Im}(Z_{yx})}{\text{Re}(Z_{yx})} \right]$$

در معادلات فوق $\text{Im}(Z_{yx})$ و $\text{Im}(Z_{xy})$ بخش موهومی و $\text{Re}(Z_{yx})$ و $\text{Re}(Z_{xy})$ بخش حقیقی امپدانس‌های الکتریکی می‌باشد. واضح است که در زمین یک بعدی $\rho_{xy} = \rho_{yx} = \rho_a$ و $|\varphi_{yx} - \varphi_{xy}| = 180^\circ$ که مقاومت ویژه ظاهری و تابع فرکانس می‌باشد. در وضعیت یک بعدی هنگام عبور از یک محیط مقاوم به یک محیط رسانا مقدار فاز از $\frac{\pi}{4}$ بالاتر می‌رود و برعکس هنگام

عبور از یک محیط رسانا به یک محیط مقاوم مقدار فاز از $\frac{\pi}{4}$ کمتر می‌شود [مرادزاده، ۱۹۹۸].

۵-۲ - ساختارهای دوبعدی

یک ساختار دوبعدی ژئوالکتریکی ساختاری است که، مقاومت ویژه الکتریکی و میدان‌ها علاوه بر تغییر نسبت به عمق در یک جهت افقی نیز به طور جانبی تغییر می‌کنند. در ساختارهای دوبعدی خواص الکتریکی به موازات یک جهت افقی ثابت است که به آن جهت امتداد^۱ می‌گویند. برای یک مدل دوبعدی معادلات درهم‌کنش امواج تخت EM برای دو وضعیت بررسی می‌شود. زمانی که مؤلفه میدان الکتریکی در جهت امتداد باشد که این وضعیت را حالت الکتریکی عرضی^۲ (TE) یا E_{\parallel} موازی (E_{\parallel}) می‌نامند. و حالتی که مؤلفه میدان الکتریکی عمود بر امتداد باشد یا مؤلفه میدان مغناطیسی در جهت امتداد باشد که این وضعیت را حالت مغناطیسی عرضی^۳ (TM) یا E_{\perp} عمودی می‌نامند. رفتار این دو نوع قطبش در برخورد با ناهمگنی‌های جانبی متفاوت می‌باشد. در حالت E_{\parallel} میدان الکتریکی در جهت امتداد می‌باشد و جریان‌های الکتریکی تمایل به شارش در جهت مؤلفه میدان الکتریکی دارند، در نتیجه آن‌ها روی سطح ناپیوستگی باری را القاء نمی‌کنند. در نتیجه اطلاعات حاصل از پاسخ E_{\parallel} توسط ناهمگنی‌های جانبی کمتر مغشوش می‌شود و آن‌ها نسبت به این ناپیوستگی‌ها حساسیت کمتری دارند. اما در مورد E_{\perp} مؤلفه میدان الکتریکی عمود بر امتداد است و جریان‌های الکتریکی را وادار به عبور در راستای عمود بر امتداد می‌کند، این میدان‌ها باعث القای بارهای سطحی می‌شوند. میدان الکتریکی ایجاد شده توسط چگالی بار سطحی به طور برداری به میدان‌های القاء شده توسط H در هر محیط اضافه می‌شود. افزایش میدان الکتریکی در نزدیکی ناپیوستگی در قسمت مقاوم باعث افزایش امپدانس و در نتیجه مقاومت ویژه ظاهری می‌شود و کاهش میدان الکتریکی در قسمت رسانا باعث کاهش مقاومت ویژه ظاهری می‌شود.

حالت E_{\parallel} به پروفیل مقاومت ویژه که مستقیماً زیر محل قرار دارد پاسخ می‌دهد یعنی این که با کاهش فرکانس و دسترسی به عمق نفوذ بیشتر، حالت E_{\parallel} اثرات ناهمگنی‌های کم عمق‌تر را نادیده می‌گیرد. اما حالت E_{\perp} حتی ناهمگنی‌های کم عمق را برای فرکانس پایین حس کرده و اثر آن هرگز از پاسخ E_{\perp} حذف نمی‌شود، این امر باعث جابه‌جایی مستقل از فرکانس منحنی مقاومت ویژه به بالا یا

1-Strike direction
2-Transverse Electric
3-Transverse Magnetic

پایین می‌شود که به این پدیده، جابه‌جایی استاتیکی^۱ می‌گویند. به همین دلیل از مد TE برای شناسایی ساختارهای عمیق و از مد TM برای شناسایی ساختارهای کم عمق استفاده می‌شود [دوبرین و ساویت، ۱۹۸۸].

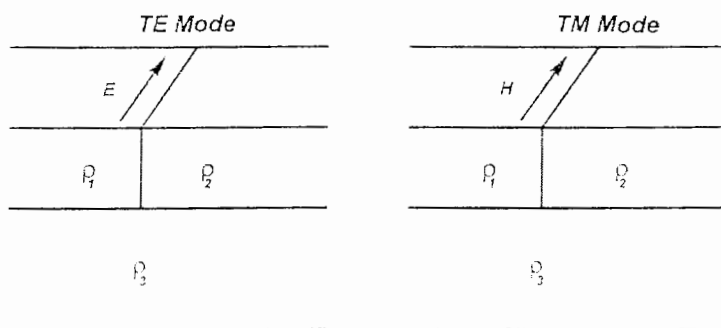
با توجه به معادله (۱-۲) مشاهده می‌شود که مؤلفه قائم H زمانی ایجاد می‌شود که $\nabla \times E$ مؤلفه قائم داشته باشد. یعنی زمانی که چگالی جریان افقی دارای تغییرات جانبی به صورت $\frac{\partial j_x}{\partial x}$ یا $\frac{\partial j_y}{\partial y}$ باشد. در این صورت میدان مغناطیسی قائمی توسط ساختار القاء می‌شود که ناشی از چشمه امواج مگنتوتلوریک نیست. با نرمالیزه کردن H_z توسط میدان مغناطیسی افقی پارامتری به نام تیپر^۲ ایجاد می‌شود که خود اطلاعات سودمندی را برای مطالعه ساختارهای زیر سطحی مهیا می‌سازد [مرادزاده، ۲۰۰۳].

۲-۶. معادلات موج الکترومغناطیسی برای هر قطبش

مدلی را مطابق شکل ۱-۲ در نظر می‌گیریم در این مدل تغییرات میدان در راستای x (جهت امتداد) ثابت است یعنی $\frac{\partial}{\partial x} = 0$.

برای حالت TE معادله موج به صورت زیر می‌باشد:

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} - j\omega\mu\sigma E_x = 0 \quad (2-28)$$



شکل ۱-۲- مدل دو بعدی و دو وضعیت قطبش TE و TM

- 1-Static shift
- 2-Tipper

و برای وضعیت TM نیز داریم:

$$\frac{\partial^2 H_x}{\partial y^2} + \frac{\partial^2 H_x}{\partial z^2} - j\omega\mu\sigma H_x = 0 \quad (29-2)$$

چگونگی محاسبه معادلات مربوط به هر قطبش در بخش اول پیوست الف ارائه می‌شود.

برای حل معادلات (28-2) و (29-2) راه حل تحلیلی وجود ندارد، بنابراین میدان‌های E و H در این معادلات دیفرانسیلی، توسط روش‌های مدل سازی عددی مانند روش اجزای محدود و تفاضل‌های محدود تقریب زده می‌شود [مرادزاده، 1998].

۲-۶-۱- تانسور امیدانس^۱

همان‌طور که قبلاً اشاره شد در مورد اکتشافات MT در یک زمین همگن یا لایه‌ای افقی رابطه بین میدان‌های الکتریکی و مغناطیسی به صورت (26-2) تعریف می‌شود. در یک زمین با ساختار دوبعدی یا سه بعدی که در طبیعت معمول تر می‌باشد، ساختار الکتریکی زیر سطحی را نمی‌توان با امیدانس اسکالر مختلط بیان کرد. برای چنین ساختارهایی رابطه بین میدان‌های الکتریکی و مغناطیسی در یک ایستگاه به صورت تانسور امیدانس ارائه می‌شود. برای یک ساختار دو بعدی تانسور امیدانس به صورت زیر می‌باشد [کنتول^۲ 1960].

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix} \begin{pmatrix} H_x \\ H_y \end{pmatrix} \quad (30-2)$$

و یا

$$E_x = Z_{xx}H_x + Z_{xy}H_y \quad (31-2)$$

$$E_y = Z_{yx}H_x + Z_{yy}H_y$$

رابطه فوق نشان می‌دهد که مؤلفه میدان الکتریکی در یک جهت معین نه تنها به مؤلفه میدان مغناطیسی عمود بر آن بستگی دارد بلکه به مؤلفه موازی با آن نیز بستگی دارد. مؤلفه‌های تانسور امیدانس مختلط و تابعی از فرکانس می‌باشند. با توجه به رابطه $E = ZH$ تانسور امیدانس (Z) به

1-Impedance tensor

2-Cantwell

صورت یک تابع تبدیل^۱ ظاهر می‌شود. در مورد ساختارهای یک بعدی $Z_{xx}=Z_{yy}=0$ و $Z_{xy}=-Z_{yx}$ می‌باشد، بنابراین تانسور امپدانس برای چنین ساختاری در غیاب نویز به صورت زیر است.

$$Z_{1-D} = \begin{pmatrix} 0 & Z \\ -Z & 0 \end{pmatrix} \quad (۳۲-۲)$$

در مورد ساختار دو بعدی چنانچه یکی از محورها به موازات امتداد ساختار باشد:

$$Z_{xx} = Z_{yy} = 0$$

$$Z_{xy} \neq Z_{yx}$$

با این شرایط تانسور امپدانس برای داده‌های فاقد نویز به این صورت است:

$$Z_{2-D} = \begin{pmatrix} 0 & Z_{xy} \\ Z_{yx} & 0 \end{pmatrix} \quad (۳۳-۲)$$

از آن جایی که در زمان برداشت ندرتاً امتداد آنومالی را می‌دانیم بنابراین محورهای برداشت در زاویه‌های دلخواه نسبت به امتداد آنومالی قرار می‌گیرند و بعداً جهت تعیین امتداد ساختار، سعی می‌شود عناصر قطری تانسور امپدانس را با دوران آن حذف کرد. البته چنانچه نویز داشته باشیم این عناصر حذف نمی‌شوند ولی می‌توان آن‌ها را به حداقل رسانید. جهت به دست آوردن محورهای موازی و عمود بر امتداد آنومالی که به آن‌ها محورهای اصلی^۲ نیز می‌گویند، محورهای برداشت را تحت زاویه θ به صورت ساعتگرد حول محور Z دوران می‌دهیم. این کار از نظر ریاضی با یک ماتریس دوران (R) به صورت زیر انجام می‌شود.

$$\begin{pmatrix} E'_x \\ E'_y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix} \quad (۳۴-۲)$$

$$E' = RE \quad (۳۵-۲)$$

همچنین در مورد H نیز به دست می‌آید:

$$H' = RH \quad (۳۶-۲)$$

1-Transfer function

2-Principle axes

که در این معادلات $R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ ماتریس دوران می باشد.

و بنابراین

$$Z'(\theta) = RZR^T \quad (37-2)$$

R^T ترانپوز R می باشد.

رابطه بین امپدانس های تانسوری و امپدانس های اصلی (Z'_{yx}, Z'_{xy}) در بخش دوم پیوست الف ارائه می شود. زاویه θ باید طوری انتخاب شود که عناصر قطری تانسور Z' صفر و یا حداقل شود.

$$Z'(\theta_0) = \begin{pmatrix} 0 & Z'_{xy} \\ Z'_{yx} & 0 \end{pmatrix} \quad (38-2)$$

به این ترتیب عناصر غیر قطری $Z'(\theta_0)$ به عنوان مؤلفه های اصلی تانسور امپدانس شناخته می شوند که با استفاده از آن ها می توان مقاومت ویژه ظاهری و فاز را به صورت زیر به دست آورد.

$$\rho'_{xy} = 0.2T|Z'_{xy}|^2 \quad \varphi'_{xy} = \tan^{-1}\left(\frac{\text{Im } Z'_{xy}}{\text{Re } Z'_{xy}}\right) \quad (39-2)$$

$$\rho'_{yx} = 0.2T|Z'_{yx}|^2 \quad \varphi'_{yx} = \tan^{-1}\left(\frac{\text{Im } Z'_{yx}}{\text{Re } Z'_{yx}}\right)$$

مرحله پایانی پردازش داده ها، مدل سازی آن ها می باشد، به این ترتیب پس از محاسبه مقادیر مقاومت ویژه و فاز در هر فرکانس، از روش های مدل سازی جهت تعیین ساختار زیر سطحی استفاده می شود. از این رو در فصل بعد به روش های مختلف مدل سازی، مزایا، معایب و مشکلات اجرایی آن ها اشاره می شود.

فصل سوم:

مدل سازی داده های ژئوفیزیکی

مدل‌سازی داده‌های ژئوفیزیکی

۳-۱- مقدمه

با توجه به این که هدف مطالعه حاضر بررسی توانمندی شبکه‌های عصبی در مدل‌سازی معکوس داده‌های مگنتوتلوریک می‌باشد، لازم است ابتدا انجام مدل‌سازی داده‌های ژئوفیزیکی، روش‌های مختلف مدل‌سازی، مزایا و معایب هر روش، مورد بررسی قرار گیرد تا ضرورت استفاده از شبکه‌های عصبی آشکار شود. از سوی دیگر برای آموزش شبکه نیاز به پاسخ مدل‌های مصنوعی که با استفاده از نرم‌افزارهای خاص تولید می‌شود، است. از این رو مباحث این فصل شامل بررسی موارد فوق و معرفی نرم‌افزار EMIXMT2D که برای تولید داده‌های مصنوعی مورد استفاده قرار می‌گیرد، می‌باشد.

۳-۲- هدف از انجام مدل‌سازی

به طور کلی هدف از مدل‌سازی، بازسازی یک مدل زمینی براساس داده‌های صحرایی می‌باشد. مدل‌سازی به دو صورت انجام می‌گیرد.

۱. مدل‌سازی فیزیکی^۱

۲. مدل‌سازی عددی^۲

در مدل‌سازی فیزیکی، یک مدل فیزیکی تهیه شده و اندازه‌گیری‌ها در مقیاس آزمایشگاهی صورت می‌گیرد. نتیجه به دست آمده با منحنی صحرایی مقایسه می‌گردد. سپس پارامترهای فیزیکی مدل آزمایشگاهی آن قدر تغییر می‌یابد تا بین این دو منحنی برازش مناسبی صورت پذیرد. با پیشرفت تکنولوژی و پیدایش کامپیوترهای سریع، امروزه این نوع مدل‌سازی کاربرد چندانی ندارد.

1-Physical modeling
2-Numerical modeling

مدل‌سازی عددی به دو نوع مدل‌سازی مستقیم یا پیشرو و مدل‌سازی معکوس تقسیم می‌شود. هر دو نوع مدل‌سازی، رابطه بین توزیع خواص فیزیکی زمین مورد مطالعه و پاسخ ژئوفیزیکی را بیان می‌کنند.

تخمین مشخصات مدل زیرسطحی زمین، با استفاده از داده‌های ژئوفیزیکی مشاهده‌ای، مدل‌سازی معکوس یا وارون نامیده می‌شود. در مدل‌سازی مستقیم از روابط ریاضی مانند معادله موج، جهت محاسبه پاسخ مدل زمینی به یک محرک فیزیکی، استفاده می‌شود.

مدل‌سازی مستقیم را می‌توان به صورت تبدیل $f = T(x)$ در نظر گرفت. که در آن f پاسخ مدل و x بردار پارامترهای مدل زیرسطحی است و T تابع تبدیل (خطی و یا غیرخطی) می‌باشد که از نظر ریاضی یک فرآیند فیزیکی مشاهده‌ای را توصیف می‌کند. روش مدل‌سازی وارون را نیز می‌توان به صورت $x' = T^{-1}(y)$ نمایش داد که x' بردار پارامترهای تخمین زده شده مدل زیر سطحی است (فضای مدل) که از بردار داده y (فضای داده) ناشی می‌شود، عملگر T^{-1} تابع تبدیل معکوس از فضای داده به فضای مدل می‌باشد. علاوه بر اهمیت انتخاب مدل (یا انتخاب T) مسائل دیگری که باید مورد توجه قرار گیرد، این است که، ممکن است T^{-1} قابل تعیین نباشد و یا داده‌های جمع آوری شده حاوی نقاط مبهم باشند مثلاً ممکن است یک سونداژ الکتریکی در آشکار کردن بخش معینی از زیر سطح ناتوان باشد. در مجموع داده‌های واقعی همیشه آلوده به نویز هستند و به همین علت بیش از یک مدل می‌تواند با داده‌های مشاهده‌ای با یک سطح خطای معین برازش یابند. به بیان دیگر وارونه سازی یکتا نیست.

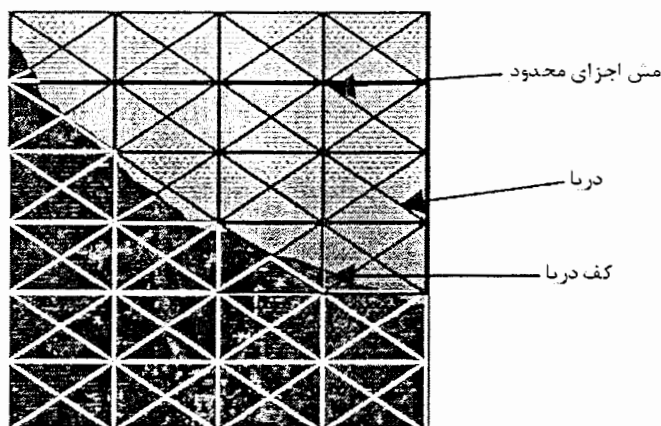
پاسخ‌های تئوری مدل‌های ژئوفیزیکی و داده‌های مشاهده‌ای با استفاده از یک الگوریتم مناسب برازش می‌شوند. تمام این الگوریتم‌ها برای کمینه کردن اختلاف داده‌های مشاهده‌ای و محاسبه شده طراحی می‌شوند. بیشتر این الگوریتم‌ها با یک حدس اولیه از پارامترهای مدل، شروع می‌شوند. در مرحله بعدی، الگوریتم، تخمینی از پارامترهای جدید را ارائه می‌دهد. این پارامترهای جدید وارد مدل تئوری می‌گردند. پاسخ مدل جدید می‌بایست برازش بهتری را با داده‌های مشاهده‌ای نتیجه دهد. اگر این اتفاق رخ دهد، وارونه سازی همگراست در غیر این صورت از روش‌های دیگری جهت دستیابی به همگرایی استفاده می‌شود. هیچ یک از روش‌های شناخته شده برای منظور فوق عاری از خطا

نمی‌باشند. در کل انتخاب یک مدل مناسب و آگاه بودن به این که چه تعداد از پارامترهای مدل می‌بایست مورد استفاده قرار گیرند و کدامیک از این پارامترها، دارای بیشترین اهمیت هستند در فرآیند مدل‌سازی بسیار مهم است. برازش خوب و مناسب بین دو منحنی شرط لازم جهت دستیابی به مدل واقعی زمین است ولی شرط کافی نیست. در واقع ممکن است در بعضی مواقع برازش خوبی بین دو منحنی وجود داشته باشد ولی مدل حاصل مدل واقعی زمین نباشد. همان طور که قبلاً اشاره شد جواب‌های به دست آمده یکتا نیستند. در واقع تعداد زیادی از جواب‌ها با یک خطای معینی می‌توانند با منحنی صحرایی برازش داشته باشند که خوشبختانه می‌توان این جواب‌ها را با اعمال قیدهایی محدود کرد [تریتل و لاینز، ۱۹۹۹].

۳-۳- مدل‌سازی مستقیم

در فرآیند مدل‌سازی مستقیم، یک مدل با پارامترهای مشخص در نظر گرفته می‌شود و پاسخ تئوری آن توسط معادلات حاکم بر مسأله محاسبه می‌شود. در تمام مسائل فیزیکی یک سری معادلات دیفرانسیل حاکم بر مسائل وجود دارد که جهت حل آن‌ها روش‌های مختلفی وجود دارد. معادلات (۲۸-۲) و (۲۹-۲) بر مسأله دو بعدی مگنتوتلوریک حاکم می‌باشند که حل آن‌ها تنها از طریق روش‌های عددی و اعمال شرایط مرزی امکان پذیر است. هدف از حل این معادلات به دست آوردن میدان‌های الکتریکی و مغناطیسی می‌باشد، با به دست آوردن این میدان‌ها می‌توان امپدانس و متعاقباً مقاومت ویژه ظاهری را محاسبه کرد. یکی از روش‌های عددی حل معادلات مذکور روش اجزای محدود می‌باشد که در این روش ابتدا منطقه مورد نظر با استفاده از خطوط قائم و افقی مش‌بندی می‌شود. هر سلول از تقاطع دو خط قائم مجاور و دو خط افقی مجاور تشکیل می‌شود که در هر گوشه از آن یک نقطه یا گره وجود دارد. هر سلول توسط دو خط قطری به چهار مثلث تقسیم می‌شود. به هر المان مثلثی یک مقاومت ویژه نسبت داده می‌شود. این مقاومت ویژه می‌تواند میانگین وزنی مقاومت ویژه تمام توده‌های حاضر در آن مثلث باشد، و یا مقاومت ویژه توده‌ای باشد که بزرگترین سطح مثلث را اشغال کرده است.

در شکل ۱-۳ یک مدل واقعی از دریا و کف آن که مش‌های اجزای محدود روی آن قرار گرفته‌اند، نشان داده شده است.



شکل ۱-۳- مدل دریا و کف آن با مش بندی اجزای محدود [راهنمای نرم‌افزار EMIXMT2D, ۱۹۹۹].

در هر المان به تعداد گره‌ها، معادله داریم. بدیهی است که با افزایش بعد مسأله، پارامترهای مسأله و متعاقباً تعداد معادلات نیز افزایش می‌یابد، از این رو دستگاه معادلات پیچیده‌تر می‌شود و دقت محاسبات کاهش می‌یابد. از سوی دیگر به دلیل خاصیت پخشی امواج مگنتوتلوریک در بسیاری از موارد شرایط مرزی عملاً برقرار نمی‌باشد بنابراین حل این معادلات با استفاده از این شرایط به طور تقریبی انجام می‌شود. پس از محاسبه پاسخ مدل این پاسخ با داده‌های مشاهده‌ای مقایسه می‌شود و سپس با توجه به محدودیت‌های زمین‌شناسی منطقه مذکور، چنانچه لازم باشد مدل اصلاح می‌شود و پاسخ تئوری مدل جدید محاسبه و با داده‌های مشاهده‌ای مقایسه می‌گردد این کار تا حصول مدلی که پاسخ آن به داده‌های مشاهده‌ای نزدیک باشد، ادامه می‌یابد. این مدل‌سازی به دلیل داشتن الگوریتم تکرار شونده و انجام کار به روش آزمون و خطا با وجود استفاده از برنامه‌های کامپیوتری، بسیار زمان‌بر می‌باشد و از طرف دیگر هیچ ضمانتی وجود ندارد که پاسخ مدل فرضی نزدیک به داده‌های مشاهده‌ای باشد و یا اصلاح مدل در جهت بهبود برازش پاسخ مدل با داده‌های مشاهده‌ای باشد.

۳-۴- مدل‌سازی معکوس

در فرآیند وارونه‌سازی، هدف محاسبه پارامترهای مدل شامل هندسه مدل، اندازه و خواص فیزیکی آن به طور مستقیم از داده‌های مشاهده‌ای می‌باشد که این امر مستلزم در نظر گرفتن یک

مدل ریاضی برای داده‌های مشاهده‌ای است. به همین دلیل این روش دارای دستگاه معادلات بسیار پیچیده‌ای می‌باشد ولی از آنجایی که هدف از مدل‌سازی داده‌های ژئوفیزیکی، به دست آوردن مدل زیر سطحی است که بهترین برازش را با داده‌های صحرائی داشته باشد. از این رو این فرآیند معمولاً با تنفیق روش‌های مدل‌سازی مستقیم و وارون صورت می‌گیرد در اغلب این روش‌ها با در نظر گرفتن یک مدل اولیه، پاسخ تئوری آن با استفاده از مدل‌سازی مستقیم محاسبه می‌شود، این پاسخ با داده‌های مشاهده‌ای مقایسه و اختلاف بین آن‌ها با استفاده از الگوریتم‌های ریاضی کمینه می‌گردد. اما همان طور که قبلاً اشاره شد، وارونه‌سازی داده‌های مشاهده‌ای مثل داده‌های حاصل از سونداژهای MT نمی‌تواند منجر به حصول یک پاسخ یکتا شود [لانگر^۱، ۱۹۳۳]. اما می‌توان با اعمال قیدهایی، تعداد مدل‌هایی را که پاسخ آن‌ها در برازش خوبی با داده‌های مشاهده‌ای می‌باشند، محدود کرد. اعمال این قیدها باید با توجه به فیزیک مسأله صورت گیرد، همچنین این قیدها نباید منجر به ناپایداری سیستم شوند. روش‌های وارونه‌سازی بر پایه نوع قیدهای مورد استفاده معمولاً به دو صورت هموار^۲ و پارامتری^۳ انجام می‌شوند.

۳-۴-۱- مدل‌سازی هموار

روش وارونه‌سازی هموار، روش ساده‌ای جهت حل مسائلی است که در آن‌ها تعداد داده‌ها کم می‌باشد. در این روش، اجازه داده می‌شود که با در نظر گرفتن قید هموارسازی، مدل تا آنجایی که ممکن است انعطاف پذیر باشد و از سوی دیگر پیچیده نباشد. در این قسمت پارامتر ناهمواری مدل، به صورت انتگرال مربع مشتق اول یا دوم پارامترهای مدل نسبت به عمق تعریف می‌شود.

$$R_1 = \int \left(\frac{dm}{dz} \right)^2 dz \quad \text{یا} \quad R_2 = \int \left(\frac{d^2m}{dz^2} \right)^2 dz \quad (۱-۳)$$

که $m(z)$ می‌تواند مقاومت ویژه یا لگاریتم مقاومت ویژه باشد. در واقع در این نوع مدل‌سازی هدف، یافتن پاسخی است که علاوه بر سازگاری با مقادیر مشاهده‌ای، دارای حداقل ناهمواری ممکن باشد.

1-Langer
2-Smooth
3-Parametric

در نمایش گسسته پارامتر ناهمواری، عملگر تفاضل به جای مشتق استفاده می‌شود.

$$R_1 = \sum_{i=2}^N (m_i - m_{i-1})^2$$

$$m(z) = m_i \quad z_{i-1} < z \leq z_i \quad i = 1, 2, \dots, N \quad (2-3)$$

$$R_2 = \sum_{i=2}^{N-1} (m_{i+1} - 2m_i + m_{i-1})^2$$

$$Z_0 = 0$$

M تعداد داده به صورت d_1, d_2, \dots, d_m است که می‌تواند مقاومت ویژه ظاهری و یا فاز در فرکانس‌های مختلف باشد. به ازای هر داده مشاهده‌ای، وزنی (σ_j) در نظر گرفته می‌شود، این وزن در واقع تخمینی از خطای موجود در داده‌ها می‌باشد. معمولاً این خطا دارای چند منبع می‌باشد که به عنوان مثال، برای روش MT می‌توان به محدود بودن دقت وسایل، اثر ناهمگنی‌های جانبی، نویز تلوریک و ... اشاره کرد.

مدل‌سازی مستقیم اجازه پیش بینی پاسخ مدل گسسته مورد نظر را به واسطه تابع $F_j[m]$ می‌دهد. جهت ارزیابی میزان برازش بین پاسخ تئوری مدل ($F_j[m]$) و داده‌های واقعی (d_j) معیار حداقل مربعات وزنی به صورت زیر تعریف می‌شود [کانستیل و همکاران، ۱۹۸۷].

$$\chi^2 = \sum_{j=1}^m (d_j - F_j[m])^2 / \sigma_j^2 \quad (3-3)$$

وزن‌دهی داده‌ها توسط ماتریس قطری شامل عناصر $1/\sigma_j^2$ ، میزان اهمیت نسبی هر داده را مشخص می‌کند. پارامترهای m در مدل‌سازی هموار باید به گونه‌ای تعیین شوند که علاوه بردستیابی به مقدار χ^2 قابل قبول پارامتر ناهمواری نیز به حداقل مقدار ممکن برسد. جهت یافتن هموارترین مدل در یک مسأله غیرخطی می‌توان به منظور ساده شدن $F_j[m]$ را به صورت یک تابع خطی از پارامترهای مدل به صورت $d = Gm$ در نظر گرفت G یک ماتریس $M \times N$ می‌باشد که عناصر آن با استفاده از تئوری مسأله مستقیم محاسبه می‌شود به طوری که تابع خطا به صورت زیر تبدیل می‌شود.

$$\chi^2 = \|Wd - WGm\| \quad (4-3)$$

W ماتریس قطری $M \times M$ به صورت $W = \text{diag}\{1/\sigma_1, 1/\sigma_2, \dots, 1/\sigma_M\}$ می‌باشد. پارامترهای ناهمواری را می‌توان به صورت یک عملگر ماتریسی به صورت $R_1 = \|\partial m\|^2$ و $R_2 = \|\partial^2 m\|^2$ نوشت. برای مورد خطی هدف کمینه کردن پارامتر ناهمواری با در نظر گرفتن این شرط که خطا در رابطه

(۳-۳) معادل با χ^2 (یک مقدار قابل قبول با توجه به خطاهای موجود در داده‌ها) باشد. یکی از روش‌های کمینه کردن تابع خطا با وجود قید هموارسازی استفاده از ضریب لاگرانژ می‌باشد [اسمیت، ۱۹۷۴]. بر اساس این روش تابعی به صورت زیر تعریف می‌شود:

$$U = \|\partial m\|^2 + \mu^{-1} \left\{ \|Wd - WGm\|^2 - \chi^2 \right\} \quad (۵-۳)$$

در این رابطه μ^{-1} ضریب لاگرانژ می‌باشد، اولین ترم در سمت راست رابطه فوق پارامتر ناهمواری و دومین ترم، خطای وزن دار شده توسط ضریب لاگرانژ می‌باشد. با صفر قرار دادن گرادیان U می‌توان پارامتر m را به صورت زیر محاسبه کرد.

$$\mu^{-1}(WG)^T WGm - \mu^{-1}(WG)^T Wd + \partial^T \partial m = 0 \quad (۶-۳)$$

$$m = [\mu \partial^T \partial + (WG)^T WG]^{-1} (WG)^T Wd \quad (۷-۳)$$

اما کمیت μ مجهول می‌باشد و می‌بایست انتخاب شود چنانچه این مقدار بزرگ انتخاب شود با توجه به تابع U مشاهده می‌شود رابطه (۷-۳) از مقدار خطا متأثر نمی‌شود و یک تابع کاملاً هموار می‌باشد و اگر این مقدار بسیار کوچک انتخاب شود ترم ناهمواری نقشی در مسأله ندارد.

حال چنانچه $F_r[m]$ به صورت یک تابع غیر خطی در نظر گرفته شود رابطه (۴-۳) به صورت زیر تبدیل می‌شود.

$$\chi^2 = \|Wd - WF[m]\| \quad (۸-۳)$$

به همان ترتیب با استفاده از ضریب لاگرانژ تابع U به صورت زیر تعریف می‌شود.

$$U = \|\partial m\|^2 + \mu^{-1} \left\{ \|Wd - WF[m]\|^2 - \chi^2 \right\} \quad (۹-۳)$$

و $\nabla_m U$ نیز به این صورت محاسبه می‌شود.

$$\mu^{-1}(WJ)^T WJm - \mu^{-1}(WJ)^T Wd + \partial^T \partial m = 0 \quad (۱۰-۳)$$

J ماتریس $M \times N$ ژاکوبین، یا گرادیان ($J = \nabla_m F$) می‌باشد. اما به دست آوردن m به سادگی حالت خطی نمی‌باشد چون در رابطه (۶-۳) G یک ماتریس ثابت می‌باشد در حالی که J به پارامتر m بستگی دارد. از این رو به جای حل یک سری معادله خطی (رابطه ۷-۳) باید یک سری معادله غیر خطی را برای به دست آوردن m حل کنیم. یکی از روش‌های حل این دستگاه معادلات غیر خطی، خطی

سازی مسأله حول یک مدل معین می‌باشد. به همین منظور ابتدا یک مدل اولیه (m_1) انتخاب می‌شود به طوری که تابع F در m_1 مشتق پذیر باشد. برای بردار کوچک Δ داریم:

$$F[m_1 + \Delta] = F[m_1] + J_1 \Delta + \varepsilon \quad \Delta = m_2 - m_1 \quad (11-3)$$

در این رابطه $J_1 = J[m_1]$ و ε جملات بعدی بسط می‌باشد که از آن‌ها صرف نظر می‌شود. چنانچه m_2 را به عنوان مدلی در نظر بگیریم که U را تحت این تقریب کمینه می‌کند. با قرار دادن این تقریب‌ها در رابطه (9-3) رابطه زیر حاصل می‌شود [کانستیل و همکاران، ۱۹۸۷].

$$U = \|\delta m_2\|^2 + \mu^{-1} \left\{ \|W(d - F[m_1] + J_1 m_1) - WJ_1 m_2\|^2 - \chi^2 \right\} \quad (12-3)$$

عبارت داخل پرانتز در ترم دوم را که نوعی بردار داده است d_1 می‌نامیم. می‌توان با استفاده از این تئوری خطی m_2 را محاسبه نمود. کمیت μ طوری انتخاب می‌شود که منجر به یک مقدار خطای قابل قبول شود.

$$m_2 = [\mu \delta^T \delta + (WJ_1)^T WJ_1]^{-1} (WJ_1)^{-1} Wd_1 \quad (13-3)$$

در این الگوریتم تکرارشونده از m_2 برای تخمین m_1, m_2, \dots, m_n که هر یک به عنوان یک مدل اولیه برای بعدی می‌باشند استفاده می‌گردد. این الگوریتم در نهایت منجر به حصول مدلی با حداقل ناهمواری و سطح خطای قابل قبول می‌شود [کانستیل و همکاران، ۱۹۸۷]. اما استفاده از روش‌های خطی‌سازی در مدل‌سازی داده‌های MT در مورد مدل‌هایی که دارای تباین شدید مقاومت ویژه می‌باشند ایجاد اشکال می‌کند. تئوری خطی‌سازی با افزایش تباین مقاومت ویژه بین واحدهای زمین‌شناسی نامعتبر می‌شود. عدم اعتبار این فرض منجر به ناپایداری این روش‌ها می‌شود.

جهت حل این معضل اخیراً راه حلی ارائه شده است [دیگروت-هلدین و کانستیل، ۲۰۰۴] در این روش از کدهای اجزای محدود ارائه شده توسط وانامیکر^۱ جهت محاسبه پاسخ تئوری مدل در مسأله مستقیم [وانامیکر و همکاران، ۱۹۸۷] و ترکیب آن با روش دلاگائو^۲ و وانامیکر برای محاسبه مشتقات جزئی $F_j[m]$ نسبت به پارامترهای مدل در ساختارهای دو بعدی استفاده می‌شود [دلاگائو و وانامیکر، ۱۹۹۶]. با وجود روش‌های بیان شده در مورد مدل‌سازی هموار این مدل‌سازی قادر به تعیین دقیق محل آنومالی و مرزهای جدایش آن‌ها نمی‌باشد از این رو از نتایج آن نمی‌توان برای حفاری

1-Wannamaker
2-Delugao

استفاده کرد [مرادزاده و عرب امیری، ۱۳۸۳]، از سوی دیگر در این مدل‌سازی به دلیل ریاضیات قضیه نمی‌توان هر محدودیتی را که از طرف فیزیک مسأله وارد می‌شود، اعمال کرد. از این رو مدل‌سازی دیگری تحت عنوان مدل‌سازی پارامتری معرفی می‌شود.

۳-۴-۲- مدل‌سازی پارامتری

(M) پارامتر مجهول (P) را در نظر بگیرید، که توسط یک تابع غیرخطی G به کمیت قابل اندازه‌گیری C مرتبط می‌شود به طوری که:

$$C=G(P,X) \quad (3-14)$$

X بردار پارامترهای معلوم مدل می‌باشد. بردار P شامل پارامترهای مجهول مدل می‌باشد به عنوان نمونه یک زمین لایه‌ای شامل پارامترهای مقاومت ویژه، ضخامت هرلایه و مقاومت ویژه نیم فضا، می‌باشد. البته مواردی وجود دارد که برخی از پارامترهای مجهول مدل از بردار P حذف و در بردار X قرار می‌گیرند. مثلاً در یک سونداژ مقاومت ویژه، مجانب‌ها می‌توانند، تخمین خوبی از مقاومت ویژه اولین لایه و یا نیم فضا باشند.

با بسط تیلور معادله مذکور و حذف اولین مرتبه آن، N معادله خطی با M مجهول حاصل می‌شود [اینمن^۱، ۱۹۷۳].

$$\Delta G = A\Delta P$$

$$\Delta G_i = G(P, X^i) - G(P^0, X^i) \quad i = 1, \dots, N$$

$$[A]_{ij} = \left. \frac{\partial G}{\partial P_j}(P, X) \right|_{\substack{X = X^i \\ P = P^0}} \quad (3-15)$$

$$[\Delta P]_j = P_j - P_j^0 \quad j = 1, \dots, M$$

بردار ΔG اختلاف بین مقاومت ویژه ظاهری اندازه گرفته شده و مقاومت ویژه ظاهری مدل اولیه P^0 در i امین فرکانس یا جدایش الکترونی در یک سونداژ مقاومت ویژه می‌باشد. A_{ij} مؤلفه‌های یک ماتریس است که عناصر آن شامل مشتق G نسبت به i امین پارامتر مجهول در P^0 و i امین فرکانس یا جدایش الکترونی در یک سونداژ مقاومت ویژه می‌شود. بردار ΔP ، اختلاف بین پارامترهای

¹I-Inman

مجهول و پارامترهای مدل اولیه می‌باشد. در این مدل‌سازی هدف، یافتن بردار پارامترهای مجهول می‌باشد به طوریکه ΔG کمینه شود.

روش‌های مختلفی جهت کمینه کردن تابع خطا (ΔG) وجود دارد از جمله می‌توان به روش‌های حداقل مربعات^۱، روش شبه نیوتنی^۲، روش گاوس - نیوتن^۳ و یا روش رگرسیون ریب^۴ (مارکوآرت^۵) اشاره کرد که این روش‌ها از نظر سرعت و دقت متفاوت می‌باشند [لاک و بارکر^۶، ۱۹۹۶؛ اینمن، ۱۹۷۵].

تخمینی که به وسیله روش حداقل مربعات از ΔP اولیه به دست می‌آید. به صورت زیر می‌باشد [اینمن، ۱۹۷۵].

$$\Delta p = (A'A)^{-1} A' \Delta G \quad (۱۶-۳)$$

در رابطه فوق A' ، ترانهاده ماتریس A می‌باشد.

بررسی‌ها نشان می‌دهد که وقتی $A'A$ به وضعیت یگانه^۷ نزدیک می‌شود [هوارل و کنارد^۸، ۱۹۷۰] اختلاف بین Δp تخمین زده شده با Δp اولیه، شروع به بزرگ شدن می‌کند. جهت جلوگیری از پیش آمدن این وضعیت روش رگرسیون ریب به این ترتیب به کار می‌رود.

$$\Delta p^* = (A'A + kI)^{-1} A' \Delta G \quad (۱۷-۳)$$

I ماتریس یکانی می‌باشد و $k \geq 0$ می‌باشد.

مقادیر ویژه ماتریس $(A'A + kI)$ برابر با $(\lambda_i^2 + k)$ می‌باشد. اضافه کردن پارامتر k به عناصر قطری ماتریس $A'A$ باعث می‌شود که مقادیر ویژه کوچک در روش حداقل مربعات توسط فاکتور k در این روش افزایش یابند. این امر باعث می‌شود که معکوس ماتریس $(A'A + kI)$ پایدارتر شود. آنچه که مهم است انتخاب مقدار مناسب برای k می‌باشد [اینمن، ۱۹۷۵]. بهترین مقدار برای k مقداری است که منجر به حداقل خطا شود. این امر می‌تواند از طریق آزمون و خطا میسر شود [هوارل و

1-least mean square

2 Quasi-Newton

3-Guass-Newton

4-Ridge Regression

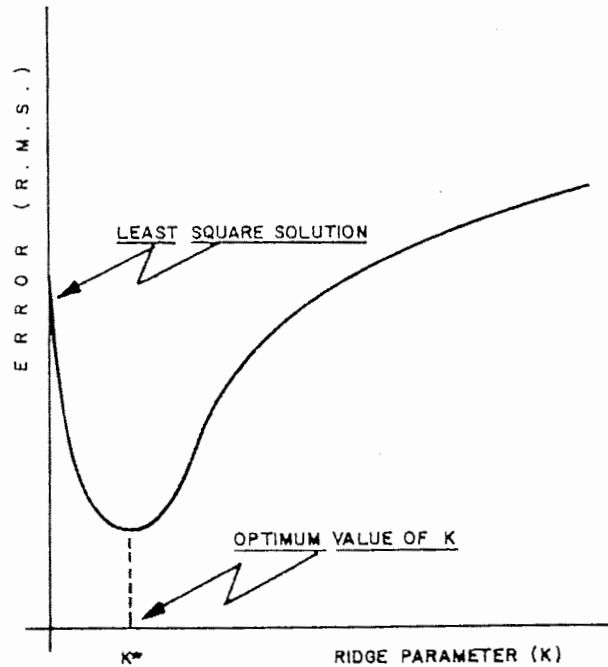
5-Marquardt

6-Loke & Barker

7-Singular

8-Hoerl & Kennard

کنارد، ۱۹۷۰b؛ کافمن و کلر^۱، ۱۹۸۱]. به این ترتیب مقدار خطا به ازای مقادیر مختلف k محاسبه و روی یک نمودار رسم می‌شود که این نمودار اثر ریدج^۲ نامیده می‌شود (شکل ۲-۳) همان طور که در شکل نیز مشاهده می‌شود در مقدار مناسب k مقدار خطا حتی کمتر از روش حداقل مربعات می‌شود [کافمن و کلر، ۱۹۸۱].



شکل ۲-۳- نمودار اثر ریدج [کافمن و کلر، ۱۹۸۱].

همان طور که مشاهده شد در تمامی روش‌های مدل‌سازی اعم از مستقیم، معکوس هموار و پارامتری، مسأله تا آن جا که ممکن است با استفاده از تقریب‌های مختلف (مثل بسط تیلور) ساده‌سازی می‌شود، از سوی دیگر در الگوریتم‌های وارونه‌سازی جهت محاسبه پاسخ تئوری مدل اولیه از روش مستقیم استفاده می‌شود و همان گونه که بیان شد با توجه به ماهیت پخشی امواج الکترومغناطیس در روش مگنتوتلوریک، استفاده از روش‌های عددی و اعمال شرایط مرزی در عمل منجر به محاسبه پاسخ تقریبی مدل می‌شود، استفاده از این پاسخ‌ها در الگوریتم‌های تکرار شونده وارونه‌سازی با تقریب‌های موجود در آن باعث افزایش خطا می‌شود. از سوی دیگر استفاده از روش‌های

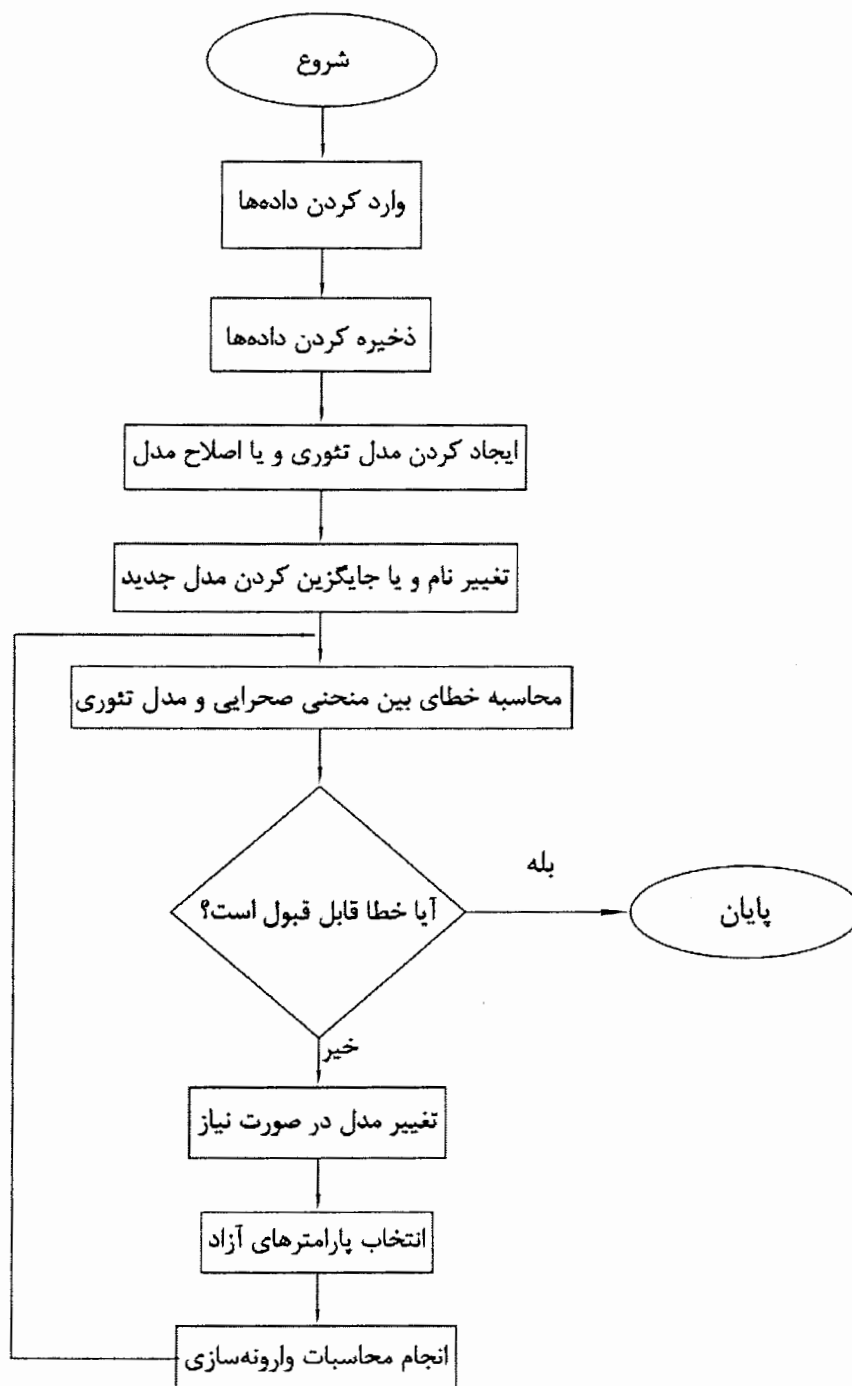
1-Kaufman & Keller
2-ridge trace

فوق در مورد مسائل مگنتوتلوریک که با داده‌های مقاومت ویژه و فاز در دو وضعیت TE و TM و بعضاً با داده‌های دامنه و فاز تیپر سر و کار داریم، ایجاد برازش به طور همزمان بین تمام این پارامترها کار بسیار مشکل و وقت‌گیری می‌باشد. از این رو محققین همواره در جستجوی روشی جدید به منظور حل مشکلات ذکر شده می‌باشند، شبکه‌های عصبی مصنوعی با توجه به توانمندی‌هایشان می‌توانند یکی از این روش‌ها باشند. اما شبکه به منظور یادگیری فرآیند وارونه‌سازی نیاز به یک سری مدل آموزشی دارد. هر مدل آموزشی شامل پارامترهای مدل به همراه پاسخ تئوری آن می‌باشد که این پاسخ توسط روش مدل‌سازی مستقیم تولید می‌شود. به همین منظور از نرم‌افزار EMIXMT2D جهت تولید مدل‌های آموزشی استفاده شده است که لازم است مختصری در مورد این نرم‌افزار و روش‌های مدل‌سازی مورد استفاده در آن توضیح دهیم.

۳-۵ - معرفی نرم افزار EMIXMT2D

این نرم افزار قابلیت مدل‌سازی دو بعدی وارون و مستقیم داده‌های مگنتوتلوریک برای مدهای TE و TM و یا هر دوی آن‌ها با هم را دارد. در این نرم‌افزار مدل‌ها شامل توده‌های چندضلعی، لایه‌ها یا ترکیبی از هر دو می‌باشند. رئوس توده، مدل را تعریف می‌کند. درکل ۱۰۰ توده (بسته یا لایه‌ای یا هر دو) و ۱۰۰۰ رأس ممکن است برای ارائه یک مدل استفاده شود. مش بندی اجزای محدود، به طور اتوماتیک پس از دریافت پارامترهای تعیین شده از سوی کاربر مانند توپوگرافی، محل ایستگاه‌ها و رسانندگی (و یا تباین رسانندگی) مدل، ایجاد می‌شود. هر سلول مستطیلی در مش به چهار المان مثلثی تقسیم می‌شود. به هر مثلث یک مقدار مقاومت ویژه بر اساس محلش نسبت به چند ضلعی‌هایی که مدل زمین را می‌سازند، نسبت داده می‌شود. نرم‌افزار فوق پاسخ تئوری مدل را با استفاده از روش اجزای محدود محاسبه می‌کند [وانامیکر و همکاران، ۱۹۸۷]. الگوریتم رگرسیون ریج اینمن [اینمن، ۱۹۷۵] از یک مدل اولیه جهت یافتن مدلی با حداقل خطا با داده‌های صحرائی استفاده می‌کند. این داده‌ها به صورت شبه مقاطع کانتور شده پریود برحسب فاصله یا پروفیل مثلاً امیدانس و مقاومت ویژه برحسب فاصله نمایش داده می‌شود [راهنمای نرم‌افزار EMIXMT2D، ۱۹۹۹].

الگوریتم وارونه‌سازی داده‌ها در این نرم افزار در فلوچارت ارائه شده در شکل ۳-۳ نشان داده می‌شود.



شکل ۳-۳- الگوریتم وارونه‌سازی نرم‌افزار EMIXMT2D

در این نرم‌افزار داده‌ها با فرمت MTX وارد و ذخیره می‌شود، سپس مدلی به عنوان مدل اولیه ایجاد و یا مدل‌های موجود اصلاح و با نام جدیدی ذخیره می‌شوند، خطای بین داده‌های صحرایی و پاسخ مدل محاسبه می‌شود و چنانچه خطا قابل قبول باشد این فرآیند پایان می‌یابد در غیر این

صورت چنانچه لازم باشد مدل اصلاح می‌شود و پارامترهای آزاد انتخاب می‌شوند و محاسبات مربوط به وارونه‌سازی انجام می‌شود و الگوریتم به مرحله محاسبه خطای بین داده‌های صحرائی و مشاهده‌ای باز می‌گردد این روند آن قدر تکرار می‌شود تا مقدار خطا به سطح قابل قبولی برسد که در آن صورت الگوریتم پایان می‌یابد.

پس از تولید مدل‌های آموزشی به وسیله این نرم‌افزار، نوبت به طراحی ساختار شبکه می‌رسد اما پیش از پرداختن به بحث چگونگی طراحی ساختار شبکه لازم است اصول و مبانی شبکه‌های عصبی، ساختار شبکه و الگوریتم‌های آموزشی آن، مورد بررسی قرار گیرند که در فصل بعد به آن پرداخته می‌شود.

فصل چهارم:

مبانی شبکه های عصبی مصنوعی

مبانی شبکه‌های عصبی مصنوعی

۴-۱- مقدمه

دیر زمانی است که بشر در پی یافتن روشی، جهت شبیه‌سازی عملکرد مغز انسان در قالب یک سیستم مصنوعی می‌باشد. این امر سبب انجام پژوهش‌های بسیاری در این زمینه شده است، که شبکه‌های عصبی مصنوعی^۱، نتیجه آن می‌باشد.

کنجکاوای جهت مطلع شدن از نحوه عملکرد مغز انسان زمانی آغاز شد که دانشمندان دریافتند مغز برای انجام محاسبات خود، از ساختاری کاملاً مغایر با ساختار کامپیوترهای متداول، برخوردار است. مغز جهت پردازش اطلاعات از ساختاری کاملاً موازی استفاده می‌کند. این ساختار موازی باعث می‌شود که حجم عظیمی از داده‌های مختلف به طور همزمان دریافت و پردازش گردد و به این ترتیب اعمال پیچیده‌ای نظیر دیدن و شنیدن میسر می‌شود [البرزی، ۱۳۸۰؛ منه‌اج، ۱۳۷۹]

۴-۲- شبکه عصبی بیولوژیکی

برای درک بهتر عملکرد سیستم‌های هوشمند، بهتر است در ابتدا به توصیف عملکرد یک نرون^۲ بپردازیم. نرون‌ها، اصلی‌ترین واحد ساختاری سیستم‌های عصبی بیولوژیکی می‌باشند. هر نرون مطابق شکل ۴-۱ از یک بدنه (سوما^۳)، اکسون^۴ و گروهی از دندریت‌ها^۵ تشکیل شده است. اکسون رشته

1-Artificial Neural Network

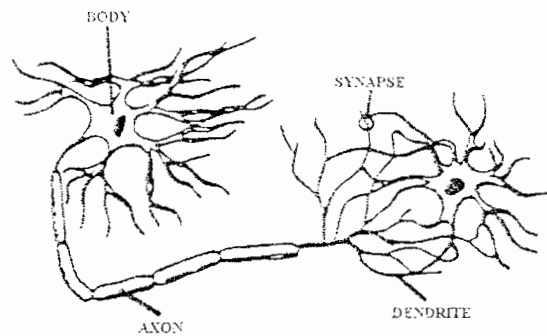
2-Neuron

3-Soma

4-Axon

5-dendrites

باریک و بلندی است که به یک سری شاخه منتهی می‌شود که انتهای این شاخه‌ها به دندیریت‌های سلول‌های دیگر وصل می‌شود. در واقع این تماس به صورت مستقیم نمی‌باشد، بلکه هر اکسون از دندیریت‌های سلول‌های مجاور توسط فاصله کوچکی موسوم به سیناپس^۱ جدا می‌شود. انتقال اطلاعات توسط سیستم عصبی بیولوژیکی ناشی از یک فرآیند الکتروشیمیایی می‌باشد. سلول عصبی در حالت سکون^۲ دارای یک عدم توازن الکتریکی، در نتیجه پتانسیلی موسوم به پتانسیل سکون می‌باشد. چنانچه سلولی تحریک گردد، قطب‌های سلول در محل تحریک موقتاً تغییر می‌کند. این تغییر ناگهانی، پتانسیل تحریک نامیده می‌شود که در طول اکسون منتشر می‌شود. البته نرون‌ها دارای این خاصیت می‌باشند که پس از تحریک، دوباره به وضعیت سکون بر می‌گردند. این پالس در انتهای اکسون به سیناپس‌ها برخورد می‌کند و باعث آزاد شدن مواد انتقال دهنده نرونی^۳ می‌شود. ماده شیمیایی آزاد شده، دارای ساختارهای مولکولی متفاوت می‌باشد که آزاد شدن آن باعث تحریک دریافت کننده‌های دندیریت نرون دیگر می‌شود. این تحریک به صورت تغییر پتانسیل محلی دندیریت‌ها بروز می‌کند که می‌تواند در جهت مثبت (محرک) یا منفی (بازدارنده) باشد. سیگنال‌های ایجاد شده وارد بدنه نرون دیگر می‌شود و چنانچه مجموع آن‌ها از مقدار آستانه نرون^۴ بیشتر باشد نرون فعال می‌شود در غیر این صورت، همان طور غیر فعال و آرام باقی می‌ماند.



شکل ۴-۱- نرون بیولوژیکی [ملسا، ۱۹۸۹].

- 1-Synapse
- 2-Resting
- 3-Neurotransmitter
- 4-Thershold of the neuron
- 5-Melsa

هر نرون دارای توانایی تطبیق‌پذیری می‌باشد. این تطبیق‌پذیری چیزی جز تقویت و تضعیف اتصالات سیناپسی نمی‌باشد. به عنوان مثال فرآیند یادگیری چهره یک شخص جدید، شامل تغییرات مختلف سیناپس‌ها می‌باشد. سیستم شبکه‌های عصبی بیولوژیکی دارای ساختاری کاملاً موازی می‌باشد که در آن تعداد زیادی واحدهای محاسباتی (نرون‌ها) به طور مشترک، انجام فعالیت را بر عهده دارند، این تقسیم کار، پیامدهای مثبتی در پی دارد. چون تعداد زیادی نرون در یک زمان درگیر فعالیت هستند، سهم هر یک از نرون‌ها، چندان حائز اهمیت نیست، بنابراین بروز مشکل در عملکرد یکی از آن‌ها، تاثیر چندانی در نتیجه کار نخواهد داشت. توان انجام فعالیت در حالی که فقط تعدادی از نرون‌ها به درستی کار می‌کنند را در علوم محاسباتی مقاوم بودن^۱ یا تحمل‌پذیری خطا می‌گویند [منهاج، ۱۳۷۹].

۳-۴- شبکه عصبی مصنوعی

دانشمندان با بررسی عملکرد شبکه عصبی بیولوژیکی و شناسایی ویژگی‌های آن، سعی در پیاده کردن این ویژگی‌ها در یک شبکه عصبی مصنوعی کردند. مهمترین ویژگی‌های شبکه عصبی بیولوژیکی، عبارتند از:

۱- تشکیل شدن از واحدهای محاسباتی ساده به نام نرون

۲- تعیین عملکرد شبکه توسط اتصالات بین نرون‌ها

۳- قابلیت یادگیری

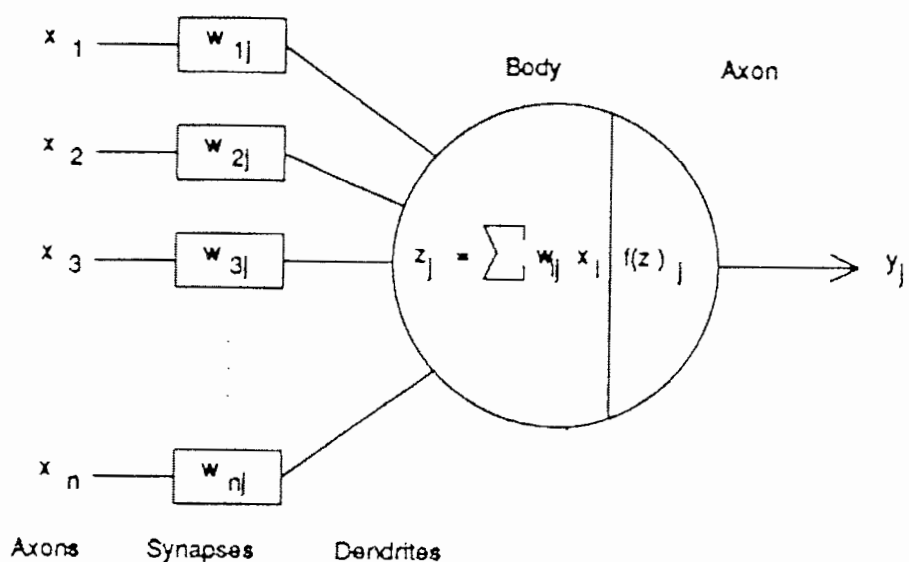
۴- برخوردار بودن از ساختاری کاملاً موازی

شکل ۲-۴ یک نرون مصنوعی را نشان می‌دهد که مدلی از رفتار نرون بیولوژیکی می‌باشد. در مقایسه آن با نرون بیولوژیکی، وزن‌ها (W) معادل سیناپس‌ها می‌باشد، که در ورودی‌ها ضرب شده و وارد بدنه نرون می‌شوند و در آن جا یک عمل جمع ساده روی آن‌ها صورت می‌گیرد. با عبور نتیجه این جمع (Z_j) از یک تابع خطی یا غیر خطی، خروجی نرون تعیین می‌شود.

$$y_j = f(Z_j) \quad (1-4)$$

$f(z_j)$ که تابع فعالیت^۱ یا محرک نامیده می‌شود. میزان فعالیت یک نرون را تعیین می‌کند. نرون مصنوعی (شکل ۲-۴) موسوم به نرون مک کلوت - پیتس^۲ می‌باشد که در آن از یک تابع آستانه‌ای باینری استفاده شده است [ملسا، ۱۹۸۹؛ پولتون و همکاران، ۱۹۹۲].

$$y_j = f(z_j) = \begin{cases} 1 & \text{if } z_j > \text{threshold} \\ 0 & \text{if } z_j < \text{threshold} \end{cases} \quad (2-4)$$



شکل ۲-۴- نرون مصنوعی [ملسا، ۱۹۸۹].

۱-۳-۴- وزن‌های ارتباطی^۳

این وزن‌ها که در این جا، نقش سیناپس‌ها را بازی می‌کنند، می‌توانند مثبت (محرک) و یا منفی (بازدارنده) باشند. شبکه‌های عصبی از اجتماع نرون‌ها تشکیل می‌شوند. وزن‌ها و توابع محرک نحوه عملکرد شبکه را تعیین می‌کنند. برخی از شبکه‌ها دارای وزن‌های قابل تطبیق^۴ می‌باشند که از

1-activation function
2-Maccolloch-Pitts
3-Connection weight
4-Adaptable weight

قوانین فراگیری^۱ جهت تطبیق دادن مقادیر وزنی استفاده می‌کنند. در نتیجه آن‌ها دارای قابلیت یادگیری می‌باشند. قابلیت یادگیری یعنی توانایی تنظیم پارامترهای شبکه (وزن‌های ارتباطی) وقتی که شبکه شرایط جدید را تجربه می‌کند. پس از آموزش شبکه توسط چند مثال، شبکه می‌تواند درمقابل یک ورودی آموزش داده نشده، پاسخ مناسب دهد. که به این امر قابلیت تعمیم^۲ می‌گویند. در یک شبکه عصبی، رفتار کلی شبکه، برآیند رفتارهای محلی نرون‌های متعدد می‌باشد. این ویژگی باعث می‌شود تا نرون‌ها در یک روند همکاری، خطاهای محلی یکدیگر را تصحیح کنند که این امر باعث افزایش قابلیت مقاوم بودن (تحمل‌پذیری خطا) سیستم می‌شود [منهاج، ۱۳۷۹؛ پولاتون، ۲۰۰۱؛ البرزی، ۱۳۸۰؛ ملسا، ۱۹۸۹].

۴-۴- تاریخچه شبکه‌های عصبی مصنوعی

پیش زمینه‌های شبکه‌های عصبی مصنوعی به اواخر قرن نوزدهم و اوایل قرن بیستم برمی‌گردد. گام‌های اولیه در این زمینه توسط روان‌شناس مشهوری به نام ویلیام جیمز^۳ در سال ۱۹۸۰ برداشته شد. وارن مک کلوت و والتر پیتس در سال ۱۹۴۳ برای اولین بار عملکرد مغز را توسط روابط ریاضی نشان دادند. دونالد‌دهب^۴ این مسیر را ادامه داد و سرانجام فرا نک روزنبلات^۵ و همکارانش در سال ۱۹۵۸ شبکه پرسپترون^۶ را معرفی کردند که این شبکه قادر به شناسایی الگوها از یکدیگر بود. ویدرووهاف^۷ در سال ۱۹۶۰ طرح روزنبلات را توسعه دادند و شبکه آدالاین^۸ با قانون فراگیری جدید را ارائه دادند. کایانلو^۹ در سال ۱۹۶۱ شبکه‌ای را عرضه کرد که شامل داده‌های متغیر با زمان بود. تئوکوهونن^{۱۰} و جیمز اندرسون^{۱۱} در سال ۱۹۷۲ شبکه‌های جدیدی را عرضه کردند که به عنوان عناصر ذخیره‌ساز عمل می‌کرد. در همین دهه استفان گروسبرگ^{۱۲} روی شبکه‌های عصبی خود

-
- 1-learning laws
 - 2-Generalization
 - 3-william James
 - 4- Donald Hebb
 - 5-Ferank Rosenblate
 - 6-Perceptron
 - 7-Widrow & Haff
 - 8-ADALINE
 - 9-Caianello
 - 10-Teo Kohonen
 - 11-James Anderson
 - 12-Stephan Grossberg

سازمانده^۱ کار کرد. جان هاپفیلد^۲ در سال ۱۹۸۲ شبکه‌های برگشتی (پسخور)^۳ را معرفی کرد و دیوید راملهارت^۴ و جیمز مک‌کلند^۵ الگوریتم پس انتشار خطا^۶ را ارائه دادند که تولدی دوباره برای شبکه‌های عصبی به شمار می‌رود. در سال‌های اخیر شبکه‌های عصبی در دو جهت تئوری و عملی توسعه بسیاری داشته است و این روند همچنان ادامه دارد. در طی این توسعه، ساختارها و روش‌های یادگیری جدید، ارائه می‌شود [منهاج، ۱۳۷۹؛ پولتون، ۲۰۰۱].

۴-۵ - کاربرد شبکه‌های عصبی

در حال حاضر شبکه‌های عصبی به طور وسیع در زمینه‌های مختلف کاربرد دارند. بخشی از این کاربردها عبارتند از: کنترل کیفیت، مخابرات، کشف جرم، بازشناسی صدا و دست خط، مسائل مالی، بیمه، امور حمل و نقل، ساخت وسایل صنعتی و پزشکی، مسایل امنیتی، مدیریت، صنایع شیمیایی، معادن و غیره. در دهه اخیر استفاده از شبکه‌های عصبی مصنوعی در زمینه ژئوفیزیک افزایش یافته است که از مهمترین کاربردهای آن، وارونه‌سازی داده‌های ژئوفیزیک می‌باشد. یک ژئوفیزیکدان یاد می‌گیرد، پاسخ زمین به اندازه‌گیری‌های فیزیکی را به دسته^۷ یا به مشخصات هدف^۸ ربط دهد. شبکه‌های عصبی نیز برای به دست آوردن رابطه بین آن‌ها، آموزش می‌بینند. بنابراین آن‌ها با استفاده از این روابط قادر به تخمین پارامترهای فیزیکی یا مشخصات هدف می‌باشند. شبکه‌های مختلفی در انواع زمینه‌های ژئوفیزیک مورد استفاده قرار گرفته‌اند. از جمله شبکه‌های به کار گرفته شده در زمینه ژئوفیزیک می‌توان به شبکه‌های زیر اشاره کرد [پولتون، ۲۰۰۱].

۱- شبکه پرسپترون چند لایه MLP

۲- شبکه RBF

۳- شبکه هاپفیلد^۹

-
- 1-Self-Organize
 - 2-John Hopfield
 - 3-feedback
 - 4-David Ramelhart
 - 5-James Mcland
 - 6-Back Propagation (BP)
 - 7-Class
 - 8-target
 - 9-Hopfield

۴- شبکه SOM¹

این شبکه‌ها در تفسیر داده‌های چاه‌پیمایی، مغناطیس، گرانی، مقاومت ویژه الکتریکی، الکترومغناطیس و لرزه‌نگاری استفاده شده‌اند [پولتون، ۲۰۰۱].

۴-۶ ساختار شبکه عصبی

تمام شبکه‌های عصبی حداقل در سه مؤلفه اصلی مشترک هستند.

۱- نرون یا گره یا واحد پردازش^۲ (PE)

۲- وزن‌های ارتباطی

۳- لایه‌های مجزایی که شامل نرون‌ها هستند و توسط وزن‌ها به هم مرتبط می‌شوند.

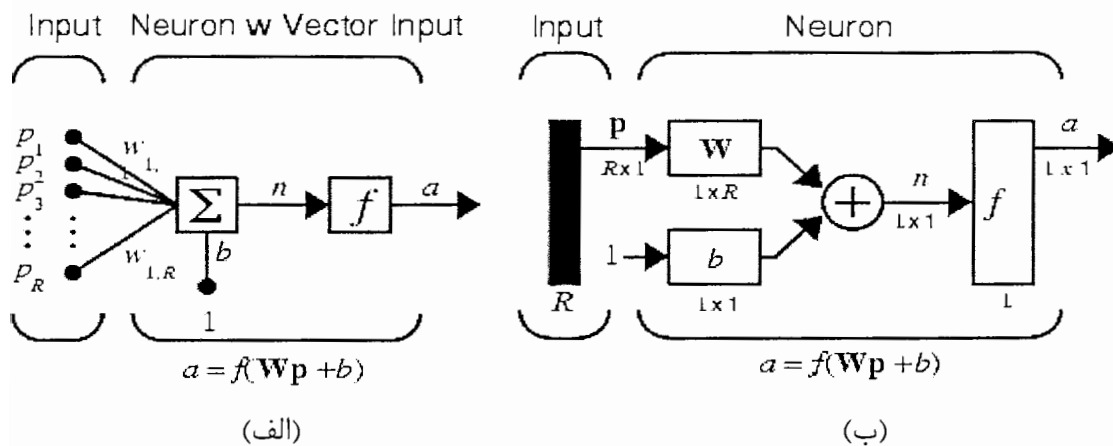
نرون‌ها به عنوان واحدهای محاسباتی پایه، اساس عملکرد شبکه‌ها را تشکیل می‌دهند. نرون‌ها شامل بردار ورودی، ماتریس وزنی، بایاس و تابع محرک می‌باشند (شکل ۴-۳). حاصل ضرب بردار ورودی (P) و ماتریس وزنی (W) و جمع آن با جمله بایاس^۳ (b) تشکیل ورودی خالص^۴ (n) را برای تابع محرک (f) می‌دهد. خروجی تابع محرک میزان فعالیت نرون را مشخص می‌کند.

$$n = \sum_{i=1}^R p_i w_{ii} + b = W \cdot P + b, \quad act_1 = f_1(n)$$

$$P = [p_1, p_2, \dots, p_R]^T \quad W = [w_{11}, w_{12}, \dots, w_{1R}] \quad (۳-۱)$$

(p_i) مؤلفه‌های بردار ورودی و (w_{ii}) مؤلفه‌های ماتریس وزن و (act_1) میزان فعالیت اولین نرون را نشان می‌دهند. لازم به ذکر است که اندیس اول در نمایش ماتریسی وزن به شماره خود نرون و اندیس دوم به مبدأ سیگنال ورودی اشاره می‌کند. در این حالت چون تنها یک نرون داریم، ماتریس وزن W، یک بردار سطری می‌باشد. بنابراین، ستون‌های ماتریس وزن بیانگر تعداد ورودی‌ها به یک نرون و سطرهای آن بیانگر تعداد نرون‌ها می‌باشد. در واقع ستون‌ها و سطرها به ترتیب ابتدا و انتهای یک خط ارتباطی را مشخص می‌کنند [منهاج، ۱۳۷۹].

1-Self Organizing Map
2-Processing Element (PE)
3-bias
4-Net input



شکل ۴-۳-۱ الف: مدل یک نرون با چند ورودی ب: فرم خلاصه شده نرون با چند ورودی [راهنمای نرم افزار matlab.

[۲۰۰۲

در مورد عملکرد شبکه، جمع وزنی ورودی‌ها به هنگام عبور از تابع محرک با آستانه تابع، مقایسه می‌شود، چنانچه این جمع وزنی از مقدار آستانه فراتر رود، نرون فعال می‌شود و اصطلاحاً آتش می‌کند و چنانچه این مقدار کمتر از آستانه باشد، نرون غیر فعال باقی می‌ماند یا اصطلاحاً آتش نمی‌کند.

$$Sum_j = \sum_{i=1}^n w_{ji} p_i \quad (۴-۴)$$

(Sum_j) مجموع وزنی ورودی به زامین نرون می‌باشد.

$$act_j = f(Sum_j) = \begin{cases} \text{آتش می کند} & \text{if } Sum_j > T \\ \text{آتش نمی کند} & \text{if } Sum_j < T \end{cases} \quad (۵-۴)$$

T مقدار آستانه تابع محرک و act_j میزان فعالیت زامین نرون می‌باشد.

با توجه به مسأله، وزن‌ها و مقدار آستانه به گونه‌ای محاسبه و تغییر می‌یابد تا هدف مسأله تأمین گردد. معمولاً مقدار آستانه به سمت چپ رابطه انتقال می‌یابد و در واقع سطح آستانه صفر در نظر گرفته می‌شود.

$$act_j = f(Sum_j) = \begin{cases} \text{آتش می کند} & \text{if } Sum_j - T > 0 \\ \text{آتش نمی کند} & \text{if } Sum_j - T < 0 \end{cases} \quad (۶-۴)$$

رابطه فوق را می‌توان به این صورت هم نوشت.

$$Sum_j = \sum_{i=1}^n w_{ji} p_i + w_{jb} \quad (7-4)$$

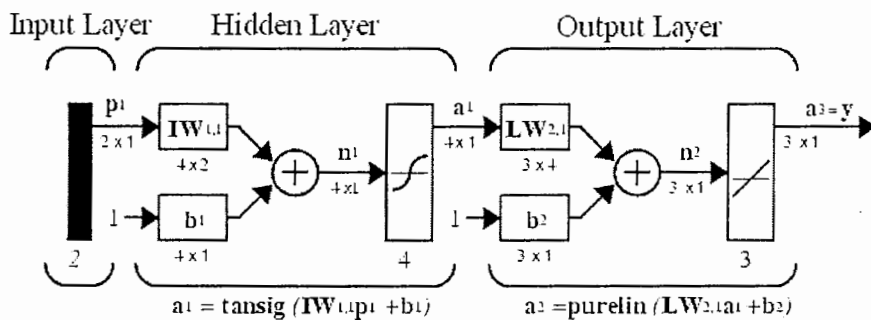
مقدار آستانه (T) را می‌توان به عنوان یک عنصر ورودی با مقدار ثابت $p_b=1$ در نظر گرفت که وزن متناظر با آن (w_{jb}) معادل با منفی مقدار آستانه قرار می‌گیرد.

$$w_{jb} = -T \quad (8-4)$$

این مقدار بیان کننده مقدار بایاس می‌باشد.

$$b = w_{jb} p_b \quad (9-4)$$

در واقع مقدار خروجی یک نرون بستگی به جمع وزنی دارد. ممکن است که این جمع در یک قسمت بهینه منحنی تابع محرک نیفتد، بنابراین جهت اطمینان از این امر جمله بایاس، برای جابه‌جا کردن آستانه به چپ و راست در طول منحنی به این جمع اضافه می‌شود [پولتون، ۲۰۰۱]. شبکه‌ها می‌توانند تک لایه و یا چند لایه (شکل ۴-۴) باشند. هر لایه شامل ماتریس وزن، جمع کننده‌ها، بردار بایاس، و تابع محرک می‌باشد. بردار ورودی (P) را می‌توان با عنوان لایه ورودی نام برد. این لایه را می‌توان شامل ماتریس W معادل با ماتریس واحد، بردار بایاس صفر و تابع محرک خطی $f(n)=n$ در نظر گرفت. در این وضعیت تعداد نرون‌های لایه ورودی^۱ برابر با تعداد عناصر بردار ورودی می‌باشد [منهاج، ۱۳۷۹]. لایه‌ای که خروجی آن خروجی نهایی شبکه است، لایه خروجی^۲ می‌نامند. تمامی لایه‌هایی که بین لایه ورودی و لایه خروجی قرار می‌گیرند، لایه پنهان^۳ نامیده می‌شوند. لازم به ذکر است که لایه‌های مختلف می‌توانند توابع محرک متفاوت داشته باشند.



شکل ۴-۴- شبکه چند لایه‌ای [اصلاح شده از راهنمای نرم‌افزار matlab، ۲۰۰۲]

- 1-Input layer
- 2-Output layer
- 3-Hidden layer

۷-۴- شبکه‌های پیشخور^۱ و پسخور

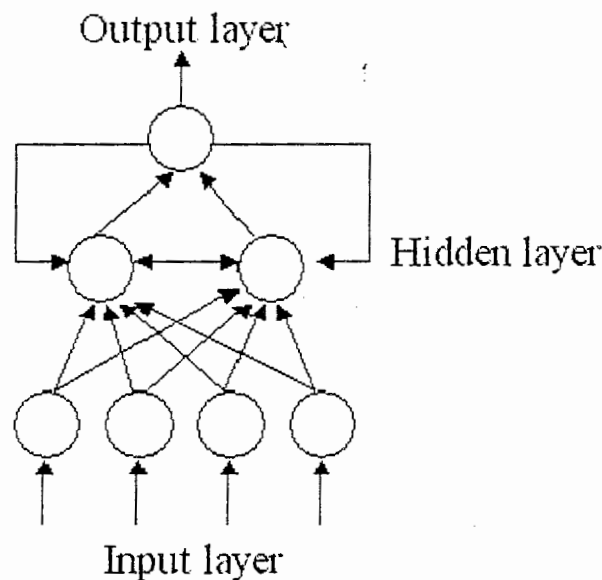
شبکه‌ها به صورت‌های مختلف دسته‌بندی می‌شوند. یکی از انواع این دسته‌بندی‌ها بر طبق جهت شارش اطلاعات می‌باشد که بر این اساس دو نوع شبکه وجود دارد.

۱-۷-۴- شبکه‌های پیشخور

در این شبکه‌ها، نرون‌ها، ورودی خود را تنها از لایه قبلی دریافت می‌کنند و خروجی خود را تنها به لایه بعدی منتقل می‌کنند. شکل ۴-۴ یک شبکه پیشخور را نشان می‌دهد [منهاج، ۱۳۷۹؛ پولتون، ۲۰۰۱؛ پولتون و همکاران، ۱۹۹۲].

۲-۷-۴- شبکه‌های پسخور یا برگشتی

مطابق شکل ۴-۵ در شبکه‌های پسخور حداقل، خروجی یک نرون، به عنوان ورودی همان نرون یا نرون‌های همان لایه و یا لایه قبلی استفاده می‌شود [منهاج، ۱۳۷۹؛ پولتون و همکاران، ۱۹۹۲].



شکل ۴-۵- شبکه پسخور یا برگشتی

۸-۴ - دسته‌بندی شبکه‌ها بر اساس روش فراگیری

نوع دیگر دسته بندی شبکه ها بر اساس نحوه فراگیری می باشد. بر این اساس شبکه ها به دو دسته تقسیم می شوند.

۱-۸-۴ - شبکه‌های دارای فراگیری با ناظر^۱

در این شبکه‌ها، جهت آموزش دادن شبکه، باید خروجی‌های مطلوب^۲ متناظر با هر الگوی ورودی^۳ تهیه شود، در این روش فراگیری، الگوهای خروجی مطلوب با خروجی محاسبه شده توسط شبکه با عنوان خروجی واقعی^۴، مقایسه می‌شود و اختلاف بین آن‌ها می‌بایست با تغییر پارامترهای شبکه (وزن‌ها و بایاس‌ها) جبران شود. نحوه تغییر پارامترهای شبکه توسط نوع الگوریتم آموزشی^۵، تعیین می‌گردد [پولتون و همکاران، ۱۹۹۲].

۲-۸-۴ - شبکه‌های دارای فراگیری بدون ناظر^۶

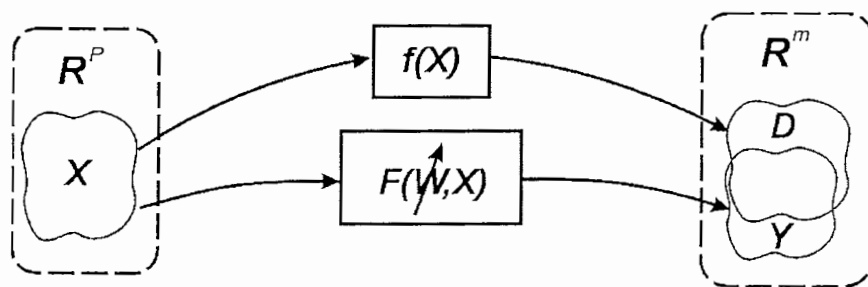
در این شبکه‌ها تنها، الگوهای ورودی جهت آموزش تهیه می‌شوند و شبکه مشخصاتی را از این الگوها استخراج می‌کند و از این مشخصات جهت دسته‌بندی الگوها استفاده می‌کند [پولتون و همکاران، ۱۹۹۲]. چگونگی ارتباط بین نرون‌ها اعم از ارتباط داخلی بین نرون‌های یک لایه با نرون‌های لایه‌های مختلف و نحوه فراگیری، نوع شبکه را مشخص می‌کند. کارایی یک شبکه توسط تعداد نرون‌ها، لایه‌ها، مقادیر وزنی، بایاس و نوع الگوریتم آموزشی مورد استفاده جهت تعیین خطاها و نحوه تغییر وزن‌ها، تعیین می‌گردد.

1-Supervised learning
2-desired output
3-input pattern
4-actual output
5-learning algorithm
6-Unsupervised learning

۹-۴- کاربرد شبکه های عصبی در ژئوفیزیک

از مهمترین کاربردهای شبکه های عصبی، دسته بندی و تقریب یا تخمین توابع^۱ با دقت مطلوب می باشد که در زمینه ژئوفیزیک نیز از این کاربردها استفاده می شود. جهت وارونه سازی داده های ژئوفیزیک معمولاً از روش فراگیری با ناظر برای تقریب توابع استفاده می شود. بنابراین ابتدا لازم است که مدل هایی برای آموزش شبکه تهیه شود. این مدل ها ممکن است با استفاده از مدل سازی مستقیم و توسط نرم افزار تهیه شود (داده های مصنوعی) و برای آموزش، استفاده گردد و یا این که جهت آموزش از داده های اندازه گیری شده (داده های صحرائی) استفاده شود. به دلیل این که برای آموزش شبکه نیاز به مجموعه داده های زیادی می باشد، غالباً از داده های مصنوعی برای آموزش استفاده می شود. انواع بسیاری از مدل ها با اشکال متفاوتی در طبیعت وجود دارد که به اندازه گیری فیزیکی صورت گرفته، پاسخ های یکسانی می دهند، چنانچه یک شبکه با انواع مختلفی از این مدل ها آموزش ببیند، دچار سردرگمی می شود پس بهتر است ابتدا این مدل ها در گروه های مختلفی قرار گیرند و برای هر گروه از این مدل ها، مثل مدل گسل ها، توده های مدفون، دایک ها و لایه های شیب دار و ... شبکه های جداگانه ای آموزش داده شود.

به هنگام آموزش یک شبکه، چنانچه X ها به عنوان ورودی و D ها به عنوان خروجی مطلوب متناظر با ورودی، به شبکه ارائه شود، شبکه سعی در به دست آوردن نگاشتی بین ورودی و خروجی می کند (شکل ۴-۶). این نگاشت به صورت $f(X)$ از فضای p بعدی ورودی (p تعداد نرون های ورودی) به فضای m بعدی خروجی (m تعداد نرون های خروجی) در نظر گرفته می شود.



شکل ۴-۶- نگاشت از فضای p - بعدی ورودی به فضای m بعدی خروجی [www.csse.monash.com, 2002].

1- function approximation

تابعی به صورت زیر:

$$f: X \rightarrow D \text{ یا } d = f(x), x \in X \subset R^p, d \in D \subset R^m \quad (10-4)$$

این تابع با استفاده از نمونه‌های آموزشی $\{X; D\}$ تعیین می‌شود. این تابع به صورت زیر توسط شبکه تقریب می‌شود.

$$F: R^p \times R^M \rightarrow R^m \text{ یا } y = F(w, x); x \in R^p, d \in R^m, w \in R^M \quad (11-4)$$

M تعداد نرون‌های لایه‌های پنهان می‌باشد.

تقریب با استفاده از روش‌های مختلف جهت کمینه کردن تابع خطای بین D و Y صورت می‌گیرد. در این روش‌ها، وزن‌ها در راستایی تغییر می‌کنند که اختلاف بین خروجی واقعی و مطلوب به حداقل برسد. بعد از مرحله آموزش و تعمیم دادن، شبکه برای استفاده آماده است و قادر به درون‌یابی و برون‌یابی داده‌های ورودی جدید می‌باشد [www.csse.monash.com, 2002]. در این تحقیق از شبکه با ناظر پرسپترون چند لایه با قانون پس انتشار خطا استفاده می‌شود، که به طور مفصل مورد بررسی قرار می‌گیرد.

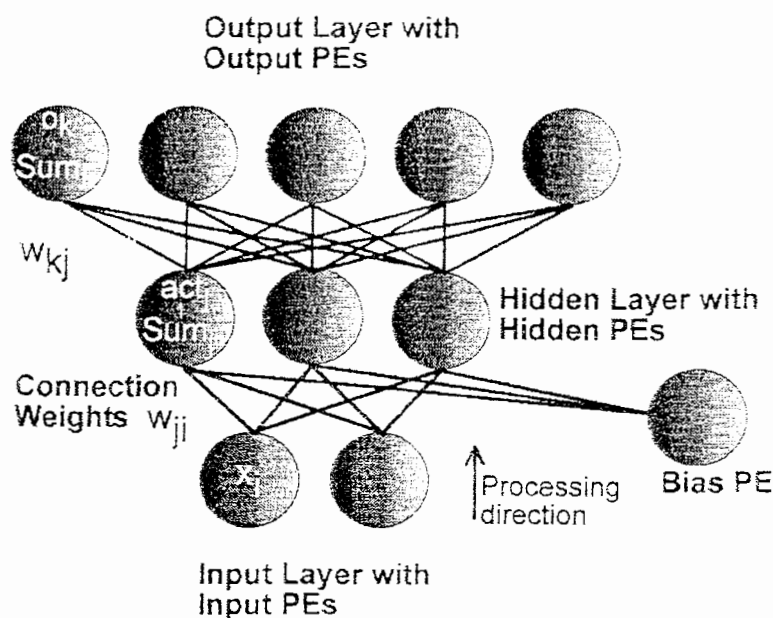
۱۰-۴-۱- پرسپترون با الگوریتم آموزشی پس انتشار خطا

شکل ۴-۷ یک پرسپترون چند لایه پیشخور که هر نرون در هر لایه به تمام نرون‌های لایه بعد متصل است، را نشان می‌دهد. معمول‌ترین طرح این شبکه شامل لایه ورودی، یک لایه پنهان و یک لایه خروجی می‌باشد. تعداد عناصر نرون‌های لایه ورودی و نرون‌های لایه خروجی، توسط مساله تعیین می‌شود و جزء پارامترهای آزاد نمی‌باشند.

مقدار فعالیت هر نرون در لایه پنهان به صورت زیر می‌باشد.

$$Sum_j = \sum_{i=1}^n w_{ji} x_i + w_{jb}, \quad act_j = f_j(Sum_j) \quad (12-4)$$

x_i : مؤلفه‌های بردار ورودی، Sum_j : مجموع وزنی ز امین نرون، w_{jb} : مقدار بایاس و act_j : میزان فعالیت ز امین نرون می‌باشد.



شکل ۴-۷- ساختار پرسپترون چند لایه پیشخور [پولتون، ۲۰۰۱]

توابع غیر خطی که معمولاً در این شبکه استفاده می‌شود. تابع سیگموئیدی و تابع تانژانت هیپربولیک می‌باشد.

$$f_j(\text{Sum}_j) = \frac{1}{1 + e^{-\text{Sum}_j}} \quad \text{تابع سیگموئیدی}$$

(۱۳-۴)

$$f_j(\text{Sum}_j) = \frac{e^{\text{Sum}_j} - e^{-\text{Sum}_j}}{e^{\text{Sum}_j} + e^{-\text{Sum}_j}} \quad \text{تابع تانژانت هیپربولیک}$$

خروجی لایه پنهان به صورت ورودی برای لایه بعد می‌باشد. سیگنال ورودی، این روند را تا دستیابی به لایه خروجی طی می‌کند. خروجی لایه بعدی که در این جا خروجی نهایی شبکه می‌باشد به صورت زیر محاسبه می‌شود.

$$\text{Sum}_k = \sum_{j=1}^m w_{kj} \text{act}_j + w_{kb} \quad (14-4)$$

$$o_k = f_k(\text{Sum}_k) \quad (15-4)$$

این شبکه از روش فراگیری با ناظر استفاده می‌کند، بنابراین الگوی خروجی مطلوب متناظر با هر الگوی ورودی، در دسترس می‌باشد. وزن‌های اولیه در شبکه به صورت تصادفی می‌باشند. خروجی که شبکه با استفاده از این وزن‌ها محاسبه می‌کند (خروجی واقعی) با خروجی مطلوب فاصله دارد و

شبکه باید این اختلاف را جبران کند. بهترین روش استفاده از تعریف خطای میانگین مربعات اصلاح شده به صورت زیر می‌باشد [پولتون، ۲۰۰۱].

$$e_p = \frac{1}{2} \sum_{k=1}^L (d_{pk} - o_{pk})^2 \quad (16-4)$$

خطای e_p برای p امین الگو به صورت مربع اختلاف بین خروجی محاسبه شده توسط شبکه o_{pk} و خروجی مطلوب d_{pk} می‌باشد. هدف از آموزش شبکه به حداقل رساندن مقدار تابع خطا با تغییر در وزن‌ها به صورت زیر می‌باشد. برای سهولت کار علامت منفی مشتق‌ها حذف شده است.

$$\frac{\partial e_p}{\partial w_{kj}} = \frac{\partial e_p}{\partial Sum_{pk}} \frac{\partial Sum_{pk}}{\partial w_{kj}} \quad (17-4)$$

$$\frac{\partial Sum_{pk}}{\partial w_{kj}} = \frac{\partial}{\partial w_{kj}} \sum_j w_{kj} act_{pj} = act_{pj} \quad (18-4)$$

متغیر جدیدی به صورت زیر تعریف می‌شود.

$$\delta_{pk} = \frac{\partial e_p}{\partial Sum_{pk}} \quad (19-4)$$

با این متغیر جدید رابطه (۱۷-۴) به صورت زیر تبدیل می‌شود.

$$\frac{\partial e_p}{\partial w_{kj}} = \delta_{pk} act_{pj} \quad (20-4)$$

معادله (۱۹-۴) را می‌توان به صورت زیر بازنویسی نمود.

$$\frac{\partial e_p}{\partial Sum_{pk}} = \frac{\partial e_p}{\partial o_{pk}} \frac{\partial o_{pk}}{\partial Sum_{pk}} \quad (21-4)$$

$$\frac{\partial o_{pk}}{\partial Sum_{pk}} = \dot{f}_k(Sum_{pk}) \quad (22-4)$$

و همچنین

$$\frac{\partial e_p}{\partial o_{pk}} = (d_{pk} - o_{pk}) \quad (23-4)$$

با جایگذاری معادلات (۲۲-۴) و (۲۳-۴) در (۱۹-۴) معادله زیر حاصل می‌شود.

$$\delta_{pk} = (d_{pk} - o_{pk}) \dot{f}_k(Sum_{pk}) \quad (24-4)$$

جهت تصحیح وزن‌ها در خطوط ارتباطی بین لایه خروجی و لایه پنهان از رابطه حاصل از جایگذاری

رابطه (۲۴-۴) در (۲۰-۴) استفاده می‌شود.

$$\Delta w_{kj} = \frac{\partial e_p}{\partial w_{kj}} = (d_{pk} - o_{pk}) f'_k (Sum_{pk}) act_{pj} \quad (25-4)$$

عبارت فوق برای نرون‌های لایه خروجی مفید است. چون در این لایه مقدار خطا برای هر نرون با توجه به سری‌های آموزشی مشخص می‌باشد ولی در لایه پنهان مقدار خروجی مطلوب برای هر نرون در دست نیست. بر این اساس مقدار خطای مربوط به هر نرون در این لایه مشخص نیست. که به این ترتیب عمل می‌شود [پولتون، ۲۰۰۱].

$$\frac{\partial e_p}{\partial w_{ji}} = \frac{1}{2} \sum_k \frac{\partial}{\partial w_{ji}} (d_{pk} - o_{pk})^2 \quad (26-4)$$

با استفاده از مشتق‌گیری زنجیره‌ای می‌توان نوشت:

$$\frac{\partial e_p}{\partial w_{ji}} = \sum_k \frac{\partial e_p}{\partial o_{pk}} \frac{\partial o_{pk}}{\partial Sum_{pk}} \frac{\partial Sum_{pk}}{\partial act_{pj}} \frac{\partial act_{pj}}{\partial Sum_{pj}} \frac{\partial Sum_{pj}}{\partial w_{ji}} \quad (27-4)$$

$$\frac{\partial e_p}{\partial w_{ji}} = \sum_k \delta_{pk} w_{kj} f'_j (Sum_{pj}) x_{pi} \quad (28-4)$$

می‌توان متغیر δ_{pj} را به این صورت تعریف کرد.

$$\delta_{pj} = f'_j (Sum_{pj}) \sum_k \delta_{pk} w_{kj} \quad (29-4)$$

در نتیجه رابطه زیر حاصل می‌شود.

$$\Delta w_{ji} = \frac{\partial e_p}{\partial w_{ji}} = \delta_{pj} x_{pi} \quad (30-4)$$

تصحیح وزن‌های بین لایه خروجی و لایه پنهان طبق رابطه (۳۱-۴) و وزن‌های بین لایه ورودی و پنهان مطابق با رابطه (۳۲-۴) صورت می‌گیرد [پولتون، ۲۰۰۱].

$$w_{kj}^{new} = w_{kj}^{old} + \alpha \delta_{pk} act_{pj} \quad (31-4)$$

$$w_{ji}^{new} = w_{ji}^{old} + \alpha \delta_{pj} x_{pi} \quad (32-4)$$

α موسوم به نرخ یادگیری^۱ می‌باشد که در واقع اندازه گام‌ها را در جهت دستیابی به حداقل خطا تنظیم می‌کند و انتخاب مناسب برای آن جهت عملکرد صحیح روش آموزشی بسیار حیاتی است.

1-Learning rate

α یک کمیت مثبت و $0 < \alpha < 2$ می‌باشد. بعد از تغییر در وزن‌ها، ورودی‌ها دوباره به کار گرفته می‌شود و این بار با وزن‌های جدید فرآیند تکرار می‌شود. به چرخه به کار گرفتن یک ورودی، محاسبه یک خروجی، محاسبه خطا و تغییر در وزن‌ها، یک تکرار^۱ گفته می‌شود. به یک بار عبور از میان سری آموزشی دوره^۲ گفته می‌شود [منیاج، ۱۳۷۹؛ ون در بین و جاتن، ۲۰۰۰].

دو وضعیت آموزشی وجود دارد، وضعیتی که در آن وزن‌ها بعد از ارائه هر الگوی آموزشی تصحیح شوند، آموزش الگو به الگو^۳، و وضعیتی که در آن وزن‌ها بعد از ارائه تمام نمونه‌های آموزشی تصحیح شوند، آموزش گروهی^۴ می‌نامند [پولتون، ۲۰۰۱، منیاج، ۱۳۷۹]. وزن‌ها با مقادیر تصادفی آغاز می‌شوند و نرخ یادگیری نیز توسط کاربر انتخاب می‌شود که در طی آموزش ثابت می‌ماند [پولتون، ۲۰۰۱].

۱۱-۴- پارامترهای شبکه

پارامترهای زیر اثر مستقیم در نحوه عملکرد شبکه دارد. کاربر می‌تواند این پارامترها را کنترل کند.

- ۱- تعداد لایه‌ها
- ۲- تعداد نرون‌های لایه پنهان
- ۳- تابع محرک
- ۴- مقدار دهی اولیه به وزن‌ها و بایاس
- ۵- نرخ یادگیری و گشتاور^۵

۱۱-۴-۱- تعداد لایه‌ها

محققین نشان دادند که یک شبکه با یک لایه پنهان که از تابع محرک سیگموئیدی استفاده می‌کند قادر به تقریب هر تابع پیوسته با تعداد معینی نرون در لایه پنهان می‌باشد. البته بیشتر

1-Iteration
2-Epoch
3-van der Bean & Jutten
4-Pattern learning
5-Batch learning
6-Momentum

تحقیقات نشان می‌دهد که یک لایه پنهان کافی است اما در بسیاری مواقع ممکن است اضافه کردن لایه پنهان دیگر باعث افزایش دقت و کاهش زمان یادگیری شود [بیشاپ^۱، ۱۹۹۵].

۴-۱۱-۲- تعداد نرون‌های پنهان

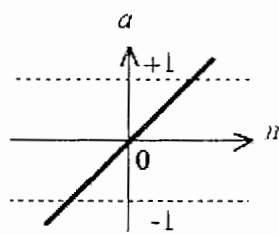
تعداد نرون‌های لایه پنهان، نقش بسیار مهمی در میزان دقت شبکه بازی می‌کند، بنابراین یافتن تعداد مناسب، اهمیت بسیاری دارد. محققین تلاش‌های بسیاری در زمینه یافتن روشی برای انتخاب بهینه تعداد نرون‌های لایه پنهان نمودند. برخی از آن‌ها واسطه هندسی تعداد نرون‌های ورودی و خروجی را به عنوان یک تعیین کننده خوب از تعداد نرون‌های لایه پنهان در شبکه‌هایی که تعداد نرون‌های خروجی در آن‌ها کمتر از ورودی می‌باشد، پیشنهاد کردند [پولتون، ۲۰۰۱]. برخی از محققین بیان کردند که برای شبکه‌ای با یک لایه پنهان، تعداد $n+1$ نرون در این لایه نیاز است که (n) تعداد نرون‌های ورودی می‌باشد [هخت - نیلسون^۲، ۱۹۸۷].

اما هیچ یک از این پیشنهادات در مورد مسائل مختلف عمومیت ندارد و تعداد بهینه نرون‌های لایه پنهان تنها از طریق آزمون و خطا قابل حصول می‌باشد [پولتون و همکاران، ۱۹۹۲].

۴-۱۱-۳- انتخاب نوع تابع محرک

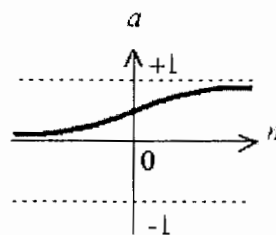
در شبکه پس انتشار خطا، هر تابع مشتق‌پذیر و پیوسته‌ای که به طور یکنواخت افزایش یابد، می‌تواند به عنوان تابع محرک استفاده شود. معمول‌ترین این توابع، سیگموئیدی (شکل ۴-۸-الف) تانژانت هیپربولیک و خطی (شکل ۴-۸-ب) می‌باشند. انتخاب نوع تابع محرک به هدف مسأله بستگی دارد. چنانچه شبکه یک مسأله دسته بندی را حل می‌کند، بهتر است از تابع سیگموئیدی یا تانژانت هیپربولیک در لایه خروجی استفاده شود. اگر هدف از حل مسأله، تقریب تابع باشد بهتر است از تابع محرک غیر خطی مانند سیگموئیدی برای نرون‌های لایه پنهان و از تابع خطی برای لایه خروجی استفاده شود [راهنمای نرم‌افزار matlab، ۲۰۰۲؛ پولتون، ۲۰۰۱].

1-Bishop
2-Hecht and Nielson



$$a = \text{purelin}(n)$$

(ب)



$$a = \text{logsig}(n)$$

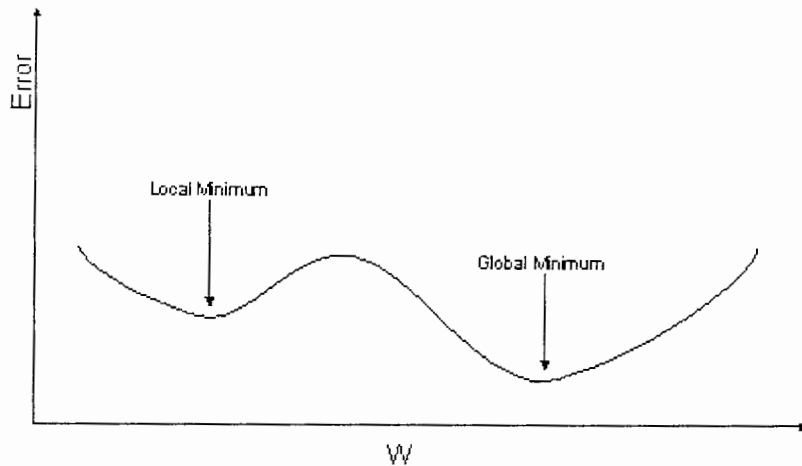
(الف)

شکل ۴-۸- الف: تابع سیگموئیدی. ب: تابع خطی [راهنمای نرم‌افزار matlab, ۲۰۰۲]

۴-۱۱-۴- مقدار دهی اولیه وزن‌ها و بایاس

انتخاب مقدار اولیه مناسب برای وزن‌ها و بایاس‌ها بسیار مهم است. چنانچه این وزن‌ها خیلی بزرگ انتخاب شوند، توابع محرک سیگموئیدی سریعاً اشباع خواهند شد و احتمال گرفتار شدن شبکه در یک کمینه محلی^۱ به جای دستیابی به کمینه سراسری^۲ را افزایش می‌دهد (شکل ۴-۹). با یک انتخاب صحیح برای مقادیر اولیه وزن‌ها، می‌توان سرعت و دقت را در شبکه افزایش داد. بهتر است این انتخاب به گونه‌ای باشد تا ورودی خالص به ژامین نرون لایه پنهان کمتر از “ البته نه خیلی کمتر از “ یک شود. این امر در صورتی محقق می‌شود که وزن‌ها از رتبه $\frac{1}{k_j}$ باشند که k_j تعداد نرون‌های ورودی است که ژامین نرون لایه پنهان را تغذیه می‌کند. مقدار مناسب اولیه برای بایاس و وزن‌ها با آزمون و خطا حاصل می‌شود. مقدار بایاس اولیه برای لایه‌های مختلف می‌تواند متفاوت باشد. حضور بایاس باعث افزایش سرعت همگرایی در شبکه می‌شود [پولتون، ۲۰۰۱].

1-Local minimum
2-Global minimum



شکل ۴-۹- کمینه‌های محلی و سراسری

۴-۱۱-۴- نرخ یادگیری و گشتاور

این دو کمیت میزان تغییر وزن‌ها را کنترل می‌کند. انتخاب نرخ یادگیری بسیار مهم است و اندازه گام‌ها را در راستای حصول به حداقل خطا تنظیم می‌کند. زمانی که تابع خطا نسبتاً تخت می‌باشد استفاده از نرخ آموزشی بزرگتر، سودمند است. استفاده از نرخ آموزشی کوچک باعث طولانی‌تر شدن زمان همگرایی می‌شود. در روش‌های آموزشی سریع‌تر که با اعمال تغییراتی در الگوریتم پس انتشار خطا حاصل شده است، نرخ یادگیری ثابت نیست و در خلال آموزش تغییر می‌کند [راهنمای نرم‌افزار matlab، ۲۰۰۲؛ پولتون، ۲۰۰۱].

گاهی اوقات شبکه به یک وضعیت پایدار دست می‌یابد ولی خروجی درستی ارائه نمی‌دهد. در واقع، شبکه در یک کمینه محلی به تله افتاده است. جهت رهایی از این کمینه محلی و سوق به سمت کمینه سراسری اضافه کردن کسری از تغییر وزن مرحله قبلی به عنوان گشتاور به معادله تنظیم ضرایب وزنی، سرعت همگرایی را نیز افزایش می‌دهد و احتمال گیر افتادن شبکه را در کمینه‌های محلی کاهش می‌دهد. گشتاور یک عدد ثابت مثبت $0 < \beta < 1$ می‌باشد اما این کمیت در برخی از الگوریتم‌های آموزشی سریع‌تر، نسبت به زمان متغیر تعریف می‌شود.

$$w_{ji}(t+1) = w_{ji}(t) + \alpha \delta_{pj} x_{pi} + \beta (w_{ji}(t) - w_{ji}(t-1)) \quad (۳۵-۴)$$

α : نرخ یادگیری

β : گشتاور

یک نرخ یادگیری کوچک می‌تواند با یک گشتاور بزرگتر جهت تسریع همگرایی استفاده شود و یک نرخ یادگیری بزرگتر معمولاً با یک گشتاور کوچکتر جهت اطمینان از پایداری شبکه می‌تواند همراه شود. مقادیر گشتاور و نرخ یادگیری اغلب توسط آزمون و خطا، انتخاب می‌شوند [پولتون، ۲۰۰۱].

۱۲-۴- پیش پردازش و پس پردازش^۱ داده‌ها

جهت افزایش کارایی شبکه غالباً از پیش و پس پردازش داده‌ها استفاده می‌شود. توابع محرک مورد استفاده در اغلب شبکه‌ها نیاز به قرار گرفتن ورودی‌هایشان در یک گستره محدود دارند. این توابع معمولاً در مقادیر کوچک و بزرگ مسطح می‌شوند و در دنباله‌هایشان دیگر به تغییر مقادیر ورودی حساس نیستند به کار بردن روش‌هایی برای قرار دادن داده‌های ورودی در یک گستره محدود مثلاً [۱، -۱] نرمالیزه کردن مقادیر یا دوباره مقیاس‌بندی کردن^۲ می‌نامند که یکی از روش‌های پیش پردازش می‌باشد [پولتون، ۲۰۰۱]. این عمل روش‌های مختلفی دارد که به عنوان نمونه می‌توان تقسیم مقادیر به بزرگ‌ترین عدد آن‌ها و یا تقسیم به نرم مقادیر ورودی اشاره کرد.

در روش‌های ژئوفیزیکی الکتریکی و الکترومغناطیسی، معمولاً با داده‌های مقاومت ویژه یا رسانندگی سر و کار داریم. گستره تغییرات این داده‌ها بسیار زیاد است، بنابراین اغلب به جای استفاده از خود داده‌ها در عملیات نرمالیزه کردن از لگاریتم آن‌ها استفاده می‌شود. داده‌های ارائه شده به نرون‌های ورودی ممکن است بر اساس نوع اندازه‌گیری دارای مقادیر بسیار متفاوت باشند و چنانچه آن‌ها در یک گستره محدود قرار نگیرند امکان دارد که شبکه نسبت به مقادیر داده‌های یکی از نرون‌های ورودی حساستر شود حتی اگر بزرگی این مقادیر مربوط به واحد اندازه‌گیری باشد. از دیگر مزایای قرار دادن ورودی‌ها در یک گستره محدود، جلوگیری از رشد زیاد وزن‌ها می‌باشد. این محدودیت، مدت زمان همگرایی شبکه و حصول کمینه خطا را کاهش می‌دهد. داده‌های خروجی نیز غالباً جهت آموزش دادن به شبکه نرمالیزه می‌شوند و قبل از آنالیز نتایج به واحدهای اصلی برگردانده می‌شوند. متعاقباً وقتی خروجی‌ها به واحدهای قبلی برگردانده می‌شود خطا نیز افزایش خواهد یافت به

1-Pre-Post processing

2-Re-scaling

ویژه وقتی که از مقیاس لگاریتمی برای آموزش استفاده می‌شود. داده‌های جدید نیز که وارد شبکه می‌شوند باید به همان طریق پیش پردازش شوند [پولتون، ۲۰۰۱].

۴-۱۴- تعمیم شبکه

شبکه‌های عصبی دارای دو خاصیت جالب توجه خلاصه سازی و تعمیم می‌باشند [تیلری و کومب^۱، ۱۹۸۷]، خلاصه سازی: توانایی استخراج مشخصه های بارز الگوی ورودی است، که نتیجه آن فشرده سازی و کد بندی کردن داده‌ها می‌باشد که این امر در لایه‌های پنهان صورت می‌گیرد. تعمیم به شبکه اجازه بازشناسی الگوهایی را که شبکه روی آن‌ها آموزش ندیده است، می‌دهد [پولتون و همکاران، ۱۹۹۲].

تعمیم شبکه، بستگی به تعداد مناسب نرون‌ها در لایه‌های پنهان و نمونه‌های آموزشی متناسب با در نظر گرفتن تعداد اتصالات وزنی دارد. چنانچه تعداد نرون‌ها در لایه پنهان به اندازه کافی نباشد. تابع تخمین زده شده روی نمونه‌های آموزشی از برازش پایینی^۲ برخوردار است، و اگر تعداد نرون‌های بیشتری مورد استفاده قرار گیرد شبکه برازش بالایی^۳ روی نمونه‌های آموزشی می‌یابد ولی با ارائه نمونه‌های جدید مقدار خطا افزایش می‌یابد. یک شبکه بزرگ نیاز به سری آموزشی بزرگی جهت ایجاد توازن بین وزن‌ها و نمونه‌های آموزشی دارد. متأسفانه مشکل است که از پیش اندازه شبکه مورد نظر را بدانیم. به همین دلیل از سایر روش های تعمیم استفاده می‌گردد.

لازم به ذکر است که شبکه‌ها در درون‌یابی عملکرد خوبی دارند لیکن در برون‌یابی چندان قوی نیستند. آن‌ها می‌توانند الگوهایی را که قبلاً دیده اند تشخیص دهند و همچنین الگوهای میانی را که ندیده‌اند در حیطه توانایی خود قرار دهند. لیکن به الگوهایی که خارج از گستره الگوهای آموزش داده شده به شبکه است، نمی‌توانند به خوبی پاسخ دهند [البرزی، ۱۳۸۰].

1-Tillery & Combs
2-underfitting
3-Overfitting

۴-۱۴-۱- منظم سازی^۱

در این روش به تابع خطا که معمولاً به صورت میانگین مربعات خطاهای شبکه روی سری آموزشی (رابطه ۴-۳۶) تعریف می‌شود.

$$F = MSE = \frac{1}{N} \sum_i^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (d_i - o_i)^2 \quad (۴-۳۶)$$

پارامتری شامل میانگین مجموع مربعات وزن‌ها و بایاس‌های شبکه (رابطه ۴-۳۷) اضافه می‌شود.

$$MSEREG = \gamma MSE + (1 - \gamma)MSW$$

$$MSW = \frac{1}{n} \sum_{j=1}^n w_j^2 \quad (۴-۳۷)$$

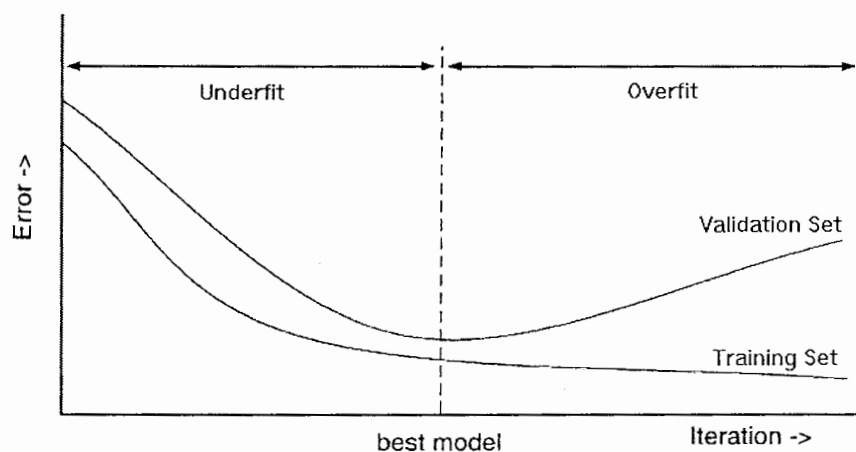
که هدف کمینه کردن این تابع خطای جدید می‌باشد. در رابطه فوق انتخاب مقدار مناسب برای γ مهم است اگر این مقدار بزرگ انتخاب شود ممکن است برازش بالا و چنانچه این مقدار کوچک انتخاب شود ممکن است برازش پایین رخ دهد [راهنمای نرم‌افزار matlab، ۲۰۰۲].

۴-۱۴-۲- توقف سریع^۲

یکی از مواردی که در آموزش شبکه با آن مواجه می‌شویم، این است که چه وقت آموزش شبکه باید متوقف شود. ممکن است ایده توقف آموزش، وقتی که شبکه به یک کمینه محلی، دست یافت یا بعد از تعداد تکرارهای معین، خوب به نظر آید. اما این امر باعث می‌شود که توزیع وزنی به دست آمده تنها، روی نمونه‌های آموزشی به خوبی پاسخ دهد [گمن^۳ و همکاران، ۱۹۹۲]، به بیان دیگر شبکه به خوبی نمونه‌های آموزشی ارائه شده را به خاطر می‌سپرد. اما وقتی نمونه‌های جدید به شبکه ارائه می‌شود. میزان خطا برای آن‌ها افزایش می‌یابد. یکی از روش‌های تعمیم شبکه، تقسیم بندی سری‌های آموزش به سه زیر شاخه می‌باشد. اولین زیر شاخه شامل سری آموزشی^۴ می‌باشد که جهت محاسبه خطا و تغییر وزن‌ها در شبکه استفاده می‌شود. دومین زیر شاخه تحت عنوان سری اعتبار سنجی^۵ از

1-Regularization
2-Early stopping
3-Geman
4-Training set
5-Validation set

داده‌هایی استفاده می‌کند که در فرآیند آموزش استفاده نمی‌شود. این سری به عنوان سری مرجع استفاده می‌شود. وقتی که خطا برای سری مرجع در یک تعداد تکرار معین افزایش یابد آموزش متوقف می‌شود و وزن‌ها و بایاس‌ها به کمینه خطا برای سری مرجع برگردانده می‌شود (شکل ۴-۱۰). زیر شاخه سوم تحت عنوان سری آزمایشی^۱ اغلب متمایز از سری اعتبار سنجی می‌باشد، این سری در آموزش استفاده نمی‌شود و تنها به عنوان معیاری برای مشخص کردن عملکرد شبکه می‌باشد، در طی فرآیندهای آزمایشی و اعتبارسنجی وزن‌ها تغییر نمی‌کنند و این داده‌ها تنها یک بار طول شبکه را طی می‌کنند [راهنمای نرم‌افزار matlab، ۲۰۰۲؛ پولتون، ۲۰۰۱؛ ون در بین و جاتن، ۲۰۰۰].



شکل ۴-۱۰- تعمیم شبکه با استفاده از سری اعتبار سنجی [اصلاح شده از ون در بین و جاتن، ۲۰۰۰].

از مسائلی که تاثیر فراوان روی عملکرد بهینه شبکه دارد نحوه ارائه داده‌های ورودی و خروجی به شبکه و تعداد نمونه‌های اعتبارسنجی و آموزشی می‌باشد. به طور ایده‌آل در یک شبکه تعداد نمونه‌های آموزشی باید بیشتر از وزن‌های ارتباطی باشد. بنابراین بردار ورودی بزرگتر، وزن‌های ارتباطی بیشتر و نمونه‌های آموزشی بیشتری نیاز دارد. به پیشنهاد پولتون، بهتر است نمونه‌های آموزش تقریباً ۱۰ برابر تعداد وزن‌ها در شبکه باشد. بررسی‌ها نشان می‌دهد که به تقریب می‌توان تعداد نمونه‌ها، برای نرون‌های خطی با توجه به سطح دقت مطلوب را محاسبه کرد [باوم و هسلر^۲، ۱۹۸۹].

$$\text{سطح دقت} = (1-e)$$

1-Test set
2-Baum & Haussler

که در رابطه فوق e مقدار خطای مطلوب شبکه می‌باشد. تعداد نمونه‌های آموزشی می‌تواند به صورت

$$\text{تعداد نمونه‌ها} = \frac{w}{e}$$

w تعداد وزن‌های ارتباطی در شبکه می‌باشد [پولتون، ۲۰۰۱؛ ون در بین و جاتن، ۲۰۰۰؛ پولتون و همکاران، ۱۹۹۲]. این امر باعث می‌شود که نحوه ارائه بردارهای ورودی مورد توجه قرار گیرد. کاهش اندازه این بردارها می‌تواند اثر بسیار مهمی روی دقت شبکه داشته باشد. البته چنانچه این کاهش به قیمت از دست دادن اطلاعات زیادی باشد، به جای افزایش دقت، کاهش دقت را در پی خواهد داشت. بنابراین موفقیت شبکه به درک ما از داده‌ها و نحوه ارائه آن‌ها به شبکه با فشرده‌ترین حالت و هنوز غنی از اطلاعات، بستگی دارد. اما چه تعداد از نمونه‌ها باید برای آموزش و چه تعداد برای تعمیم استفاده شوند. برخی از محققین دریافتند که چنانچه شبکه روی حدود ۶۰٪ از الگوهای ورودی آموزش ببینند می‌تواند با موفقیت تعمیم را روی داده‌های باقی‌مانده انجام دهد [سولی و همکاران^۱، ۱۹۸۷]. با توجه به مطالب ارائه شده در این فصل، مبنی بر آشنایی با شبکه و معرفی اصول و مبانی طراحی آن، نوبت به طراحی شبکه‌های مورد نظر می‌رسد که در فصل بعد به چگونگی طراحی آن‌ها پرداخته می‌شود.

فصل پنجم:

وارونه‌سازی داده‌های مگنتوتلوریک با
استفاده از شبکه‌های عصبی مصنوعی

وارونه‌سازی داده‌های مگنتوتلوریک با استفاده از شبکه‌های عصبی مصنوعی

۵-۱- مقدمه

همان طور که قبلاً بیان شد، شبکه‌های عصبی مصنوعی مدل‌های ساده‌ای هستند که سعی در شبیه‌سازی عملکرد نرون‌های مغز انسان، دارند. شبکه‌ها می‌توانند رابطه بین پارامترهای خروجی و ورودی به شبکه را شبیه‌سازی کنند، این رابطه ممکن است، توابع ریاضی از داده‌های مصنوعی تولید شده یا داده‌های اندازه‌گیری شده باشد. در زمینه وارونه‌سازی داده‌های ژئوفیزیکی شبکه‌های عصبی می‌توانند تخمینی از پارامترهای مدل زمینی مانند ضخامت و مقاومت ویژه لایه‌ها در یک مدل یک بعدی را، بعد از آموزش ارائه دهند.

تولید سری‌های آموزشی می‌تواند زمان‌بر باشد، در ضمن تعداد مدل‌های آموزشی نیز ممکن است زیاد باشد. وارونه‌سازی توسط شبکه‌های عصبی می‌تواند به دقت وارونه‌سازی حداقل مربعات و از همه مهمتر، سریعتر باشد [پولتون، ۲۰۰۱]، به این ترتیب که شبکه یک بار برای کلاس خاصی آموزش می‌بیند و چنانچه به خوبی آموزش ببیند و تعمیم یابد با ارائه داده‌های مشاهده‌ای در همان کلاس، قادر به تخمین پارامترهای مدل با دقت قابل قبولی در مدت زمان بسیار کوتاه خواهد شد. همچنین شبکه‌های عصبی می‌توانند جهت یادگیری مدل‌سازی پیشرو آموزش داده شوند و وقتی که این شبکه‌ها با روش وارونه‌سازی حداقل مربعات ترکیب شوند می‌توانند منجر به تسریع در وارونه‌سازی داده‌ها شوند [پولتون، ۲۰۰۱].

در این تحقیق از شبکه عصبی به منظور شبیه‌سازی رابطه بین دوسری پارامتر استفاده شده است پارامترهای ورودی در این شبکه پاسخ مدل زمینی به یک محرک فیزیکی و پارامترهای خروجی،

همان پارامترهای مدل می باشند، در روش مگنتوتلوریک، محرک، میدان‌های مغناطیسی ایجاد شده توسط چشمه‌های طبیعی می‌باشد و پاسخ مدل نیز میدان‌های الکتریکی و مغناطیسی القاء شده می‌باشد که با یک سری عملیات ریاضی به صورت مقاومت ویژه ظاهری و فاز امپدانس در هر فرکانس و برای هر دو مد TE و TM محاسبه می‌گردند.

۵-۲- انتخاب شبکه مناسب

اولین گام در به کارگرفتن شبکه‌های عصبی برای یک هدف خاص، انتخاب شبکه مناسب می‌باشد. هدف از وارونه‌سازی داده‌ها، تخمین یک تابع می‌باشد همان طور که می‌دانید پاسخ مدل زیر سطحی تابعی از پارامترهای مدل می‌باشد.

$$f = T(x)$$

هدف از وارونه‌سازی داده‌ها به دست آوردن پارامترهای مدل می‌باشد بنابراین

$$x = T^{-1}(f)$$

هر دو شبکه RBF و MLP شبکه‌هایی هستند که به طور موفقیت آمیز به منظور تخمین تابع استفاده شده‌اند زمانی که حل مسأله وارون معادل با تخمین یک نگاشت وارون است منطقی است که از این شبکه‌ها استفاده گردد [ون در بین و جاتن، ۲۰۰۰].

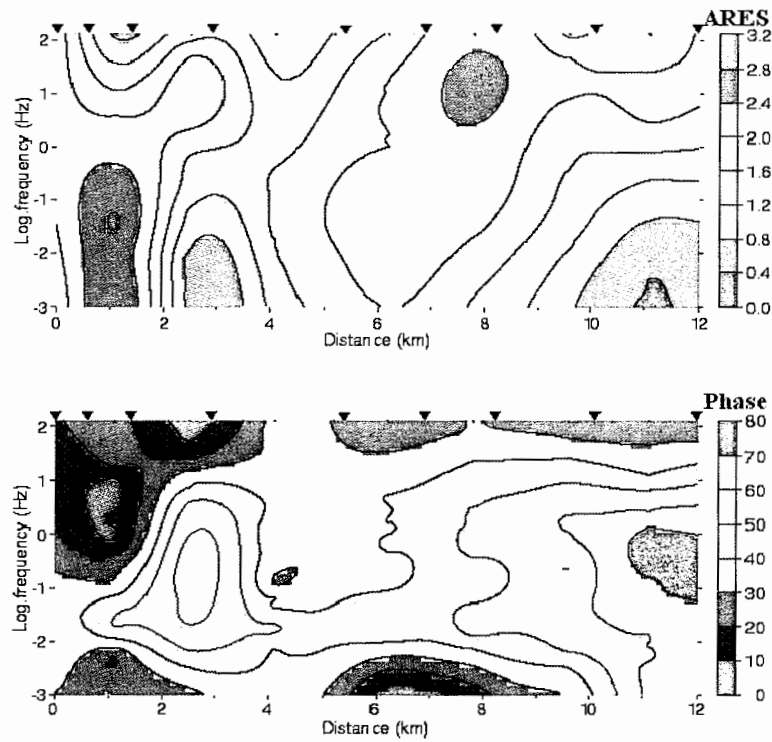
این دوشبکه از تخمین‌گرهای رایج می‌باشند که قادر به تولید نتایج یکسانی می‌باشند. با این وجود ممکن است یکی از آن‌ها نتایج مناسب‌تری در زمینه یک کاربرد خاص نسبت به نوع دیگر بدهد. بنابراین چنانچه هیچ تجربه یا راهنمایی در مورد انتخاب یکی از این دوشبکه موجود نباشد بهتر است هر دو، مورد آزمایش قرار گیرد [ون در بین و جاتن، ۲۰۰۰]. باتوجه به مقالات موجود، در این تحقیق از شبکه MLP با روش آموزشی پس انتشار خطا استفاده شده است.

۵-۳- انتخاب کلاس مدل

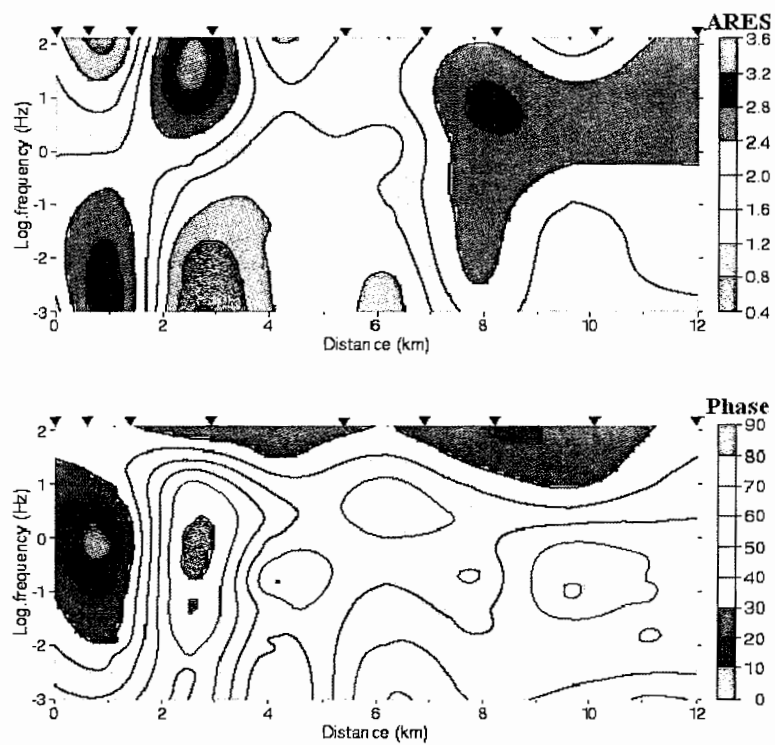
کلاس مدل بر اساس داده‌های صحرایی در دسترس انتخاب شد. این داده‌ها توسط دکتر مرادزاده در سال ۱۹۹۷ در استرالیای جنوبی در یک پروفیل عمود بر رشته کوه‌های فلیندرز^۱ و به طول ۱۲ کیلومتر برداشت شد. برداشت داده‌ها در ۹ ایستگاه و ۲۱ فرکانس صورت گرفت. شبه مقاطع مقاومت ویژه و فاز داده‌های صحرایی برای دو مد TE و TM (شکل‌های ۵-۱ و ۵-۲) با استفاده از نرم‌افزار گرافیکی GMT^۲ رسم شد [www.gmt.soest.hawaii.edu, 2004] که با توجه به شبه مقاطع رسم شده، به وضوح حضور دو آنومالی هادی مشاهده می‌شود [مرادزاده، ۱۹۹۸]. به منظور ساده کردن مسأله، زمینه آنومالی‌ها همگن در نظر گرفته شد. از این رو کلاس مدل، دو آنومالی هادی مدفون در یک نیم فضای همگن انتخاب شد. برای ساده سازی و کاهش تعداد پارامترهایی که باید تخمین زده شود شکل آنومالی‌ها به صورت مستطیل تقریب شد، شکل آنومالی‌های واقعی، هر چه باشد، می‌توان آن‌ها را در یک مستطیل محاط، یا بر یک مستطیل محیط کرد، این ساده سازی به عنوان نقطه شروع در بررسی توانمندی‌های شبکه عصبی در وارونه‌سازی داده‌های مگنتوتلوریک، مناسب به نظر می‌آید. شاید دقت شبکه عصبی با در نظر گرفتن شکل آنومالی‌ها به صورت پیچیده‌تر مانند بیضی بهبود یابد چون این مدل به مدل واقعی زیر سطحی نزدیک‌تر می‌شد. ولی تولید نمونه‌های آموزشی با این شکل هندسی با استفاده از نرم‌افزار EMIXMT2D بسیار مشکل می‌باشد. به دلیل این که این نرم‌افزار اشکال هندسی مختلف را تنها با دریافت مختصات رئوس، تولید می‌کند. در نتیجه برای تولید شکلی که بتوان به تقریب آن را بیضی در نظر گرفت نیاز به در نظر گرفتن رئوس زیادی می‌باشد، که این کار برای تولید تعداد زیادی نمونه آموزشی در دو مد، بسیار وقت‌گیر می‌باشد.

1-Flinders Ranges

2-Generic Mapping Tools



شکل ۵-۱- شبه مقاطع مقاومت ویژه (بالا) و فاز (پایین) برای داده‌های صحرائی مد TE، مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد. [مرادزاده، ۱۹۹۸]



شکل ۵-۲- شبه مقاطع مقاومت ویژه (بالا) و فاز (پایین) برای داده‌های صحرائی مد TM، مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد. [مرادزاده، ۱۹۹۸]

۵-۴ - تعیین پارامترهای ورودی و خروجی شبکه

در طراحی یک شبکه پارامترهای ورودی و خروجی شبکه، جزء پارامترهای آزاد نیستند و توسط مسأله تعیین می‌شوند. پارامترهای ورودی شبکه، پاسخ زمین به امواج مگنتوتلوریک در حوزه فرکانس می‌باشد، که شامل مقاومت ویژه ظاهری و فازامپدانس برای دو مد در ایستگاه‌های مختلف و فرکانس‌های مورد نظر می‌باشد. پارامترهای خروجی شبکه نیز همان پارامترهای مدل می‌باشد (شکل ۵-۳) که عبارتند از:

نسبت مقاومت ویژه زمین به مقاومت ویژه آنومالی می‌باشد که برای سادگی مقاومت ویژه آنومالی‌ها ثابت و برابر با $10\Omega m$ فرض شده است.

x_1, \bar{x}_1 : مختصات مرکز آنومالی سمت چپ

d_1 : نصف طول آنومالی سمت چپ

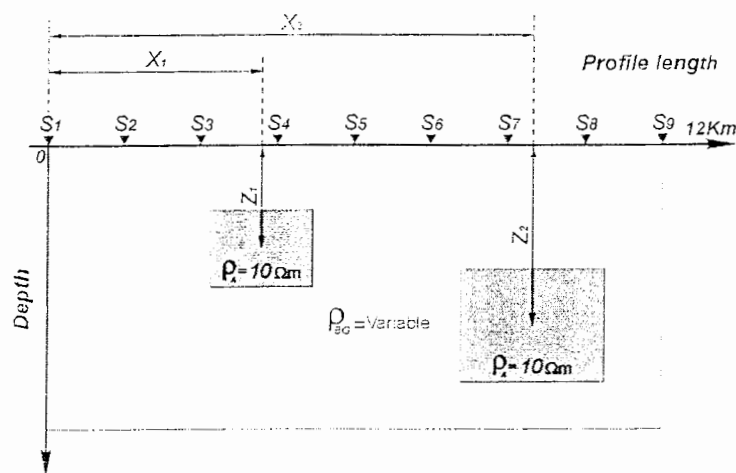
d_2 : نصف عرض آنومالی سمت چپ

x_2, \bar{x}_2 : مختصات مرکز آنومالی سمت راست

d_3 : نصف طول آنومالی سمت راست

d_4 : نصف عرض آنومالی سمت راست

به این ترتیب شبکه باید جهت تخمین ۹ پارامتر مذکور آموزش ببیند.



شکل ۵-۳- کلاس مدل انتخاب شده (دو آنومالی هادی مدفون در یک نیم فضای همگن) مثلث‌های وارون محل ایستگاه‌های اندازه‌گیری را نشان می‌دهد.

۵-۵ - تولید نمونه های آموزشی

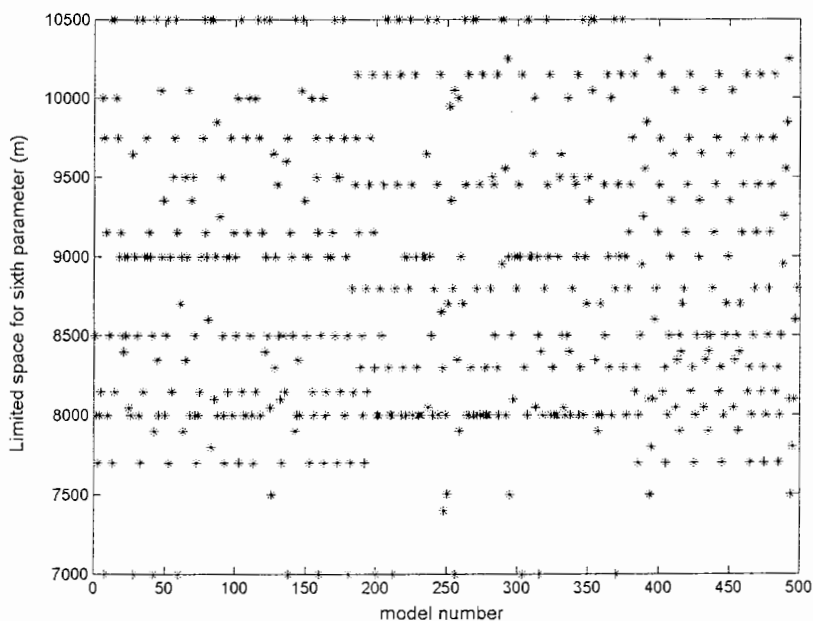
یک نمونه آموزشی از دو بردار ورودی و خروجی تشکیل می شود. بردار ورودی شامل داده های مصنوعی تولید شده در ایستگاه ها و فرکانس های مختلف و بردار خروجی شامل پارامترهای مدل می باشد. تولید نمونه های آموزشی با توجه به خروجی های شبکه (پارامترهای مدل) صورت گرفته است. با توجه به حضور ۹ پارامتر خروجی، وزن ها و بایاس های شبکه باید در خلال فرآیند آموزش به گونه ای تغییر یابند و اصلاح شوند که مجموع خطای پارامترهای خروجی شبکه کاهش یابد. بنابراین واضح است که آموزش دادن شبکه های جداگانه برای هر پارامتر خروجی غالباً منجر به نتیجه بهتری نسبت به آموزش یک شبکه واحد برای تخمین چندین پارامتر می شود [مانری و همکاران^۱، ۱۹۹۶]. این نکته در مورد تخمین پارامتر C تجربه شده است که به آن پرداخته خواهد شد.

وجود ۹ پارامتر خروجی در شبکه معادل با در نظر گرفتن یک فضای ۹ بعدی می باشد بنابراین پارامترهای مدل های تولید شده باید تقریباً این فضای ۹ بعدی را پوشش دهند. جهت تولید نمونه های آموزشی، برای هر پارامتر خروجی با تعریف کران های بالا و پایین یک محدوده معینی در نظر گرفته شده است. هرچه گام های حرکت در محدوده های تعریف شده کوچکتر باشد، داده های بیشتری جهت تقریب رابطه خروجی بر حسب ورودی فراهم می گردد. از آن جایی که در ابتدا از میزان حساسیت شبکه به داده ها اطلاعی در دست نداریم لذا پوشش فضای مدل در مورد تمام پارامترها به طور یکنواخت صورت می گیرد. نکته مهم تر در تولید نمونه های آموزشی، حفظ تباین خوب بین مقاومت ویژه زمينه و آنومالی ها با توجه به ابعاد آنومالی ها، عمق آن ها و فرکانس های اندازه گیری می باشد. فرض کنید در تولید نمونه های آموزشی، پوشش خوب فضای مدل مد نظر قرار گرفته است اما علیرغم این موضوع، شبکه ممکن است پس از آموزش به خوبی پاسخ ندهد. علت این امر عدم رعایت تباین مقاومت ویژه آنومالی ها و زمينه با توجه به عمق آنومالی ها، پریود یا فرکانس اندازه گیری و ابعاد آنومالی ها می باشد. مثلاً چنانچه در چندین نمونه آموزشی تولید شده که در آن ها آنومالی های کوچکی در قسمت های عمیق زمین قرار گرفته باشند. تباین مقاومت ویژه بسیار ضعیف باشد، داده های ثبت شده با توجه به پریودهای اندازه گیری ممکن است حضور آنومالی را در این نمونه ها نشان ندهد. به

1-Manry et al

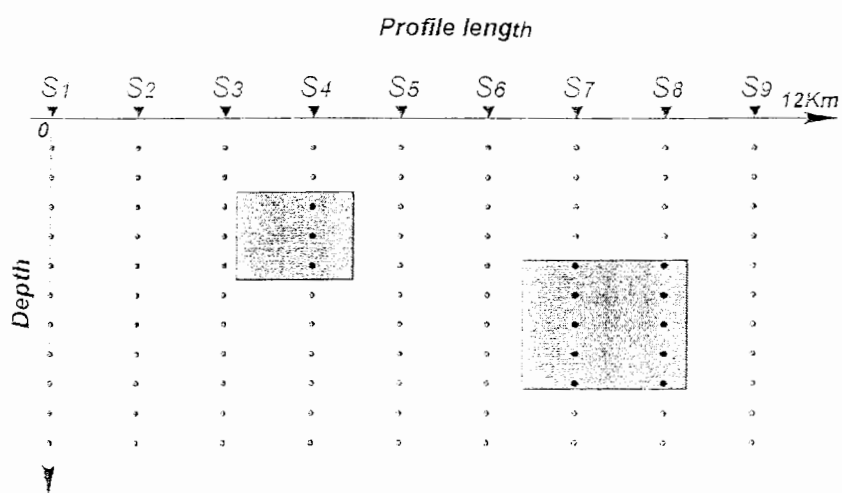
بیان دیگر مثل این است که شبکه به ازای چندین الگوی ورودی یکسان، الگوهای خروجی کاملاً متفاوت دریافت کند این امر باعث می‌شود که به شبکه یک سری نمونه‌های آموزشی اضافی ارائه شود که هیچ نقشی جز افزایش خطا در شبکه، طی فرآیند یادگیری ندارند. این نکته به وضوح پس از آموزش شبکه مشاهده می‌شود که خطا روی مدل‌هایی که تباین مقاومت ویژه ضعیف‌تری داشته‌اند و آنومالی‌ها از ابعاد کوچکتری نیز برخوردار بودند، افزایش یافته است. بنابراین شرط اصلی برای دستیابی به خطای کمتر در اختیار داشتن تعداد بیشتری از داده‌های مناسب می‌باشد. به هر اندازه تعداد نمونه‌های آموزشی با خروجی متفاوت، اما ورودی نزدیک به هم در بین داده‌ها بیشتر باشد، شبکه دچار خطای بیشتری خواهد شد.

بنابراین خطای شبکه در مناطقی از فضای مدل که چگالتر است و تباین مقاومت ویژه بهتری دارد، کاهش می‌یابد. این امر نشان می‌دهد که سری داده‌های آموزشی نقش بسیار مهمی را در نحوه عملکرد شبکه‌های عصبی بازی می‌کند [دای و همکاران^۱، ۱۹۹۹]. شکل ۴-۵ پوشش محدوده تعیین شده برای پارامتر ششم (x_2) را نشان می‌دهد. همان‌طور که در این شکل مشاهده می‌شود محدوده تعریف شده برای این پارامتر از ۷۰۰۰ تا ۱۰۵۰۰ متر می‌باشد که با ۴۹۸ مدل پوشش یافته است.



شکل ۴-۵- پوشش محدوده تعریف شده برای پارامتر ششم (x_2)

باتوجه به برداشت داده‌ها در دو جهت برای ثبت مدل دو بعدی زیر سطحی تعداد داده‌های ورودی به شبکه معادل با $N = N_p \cdot N_T \cdot N_r$ می‌باشد در این رابطه N_r تعداد مؤلفه‌های مورد استفاده در وارونه‌سازی می‌باشد مثلاً در ارائه یک تصویر کامل از زمین $N_r = 4$ معادل با مقاومت ویژه ظاهری و فاز در هر دو مد می‌باشد. N_T تعداد پریودها یا فرکانس‌ها و N_S تعداد ایستگاه‌های برداشت می‌باشد. با این وجود چنانچه داده‌های مصنوعی، همانند داده‌های صحرائی در ۹ ایستگاه و ۲۱ فرکانس تولید می‌شد تعداد ورودی‌های شبکه افزایش می‌یافت، در نتیجه اندازه شبکه بزرگ می‌شد و آموزش شبکه با امکانات کامپیوتری در دسترس دچار مشکل می‌شد از این رو به منظور کاهش تعداد ورودی‌های شبکه و کاهش اندازه آن، تعداد ۱۱ فرکانس از بین فرکانس‌های موجود انتخاب شد. تولید داده‌های مصنوعی با استفاده از مدل‌سازی پیشرو و با نرم‌افزار EMIXMT2D, VERSION:4 صورت گرفت. داده‌ها به صورت مقاومت ویژه و فاز در ۹ ایستگاه به فواصل ۱۵۰۰ متر و ۱۱ فرکانس ($0.171875 Hz$, $0.16641 Hz$, $0.1830 Hz$, $0.13809 Hz$, $0.1465 Hz$, $0.106 Hz$, $0.1215 Hz$, $0.122 Hz$, $0.1073 Hz$, $0.1031 Hz$, $0.1018 Hz$) در دو مد TE و TM برای هر مدل تولید شد. جهت ارائه یک تصویر کامل از مدل زیرسطحی در هر مد، دو پارامتر مقاومت ویژه ظاهری و فاز امپدانس برای 11×9 نقطه محاسبه شد، در نتیجه به ازای هر مدل در هر دو مد جمعاً ۳۹۶ داده تولید شد (شکل ۵-۵).



شکل ۵-۵- دو توده آنومالی مدفون در یک نیم فضای همگن همراه با فرکانس‌ها (پریودها) و ایستگاه‌های

اندازه‌گیری

مدل‌های آموزشی با حرکت افقی و قائم آنومالی‌ها نسبت به یکدیگر و تغییر همزمان ابعاد و مقاومت ویژه زمینه تولید شدند.

۵-۶- طراحی شبکه عصبی

پس از تولید نمونه‌های آموزشی و انتخاب نوع شبکه نوبت به طراحی ساختار شبکه عصبی و آموزش آن می‌رسد. طراحی و آموزش یک شبکه بیشتر یک هنر است تا یک علم [ون در بین و جاتن، ۲۰۰۰]. با توجه به این‌که ساختار یک شبکه از مهمترین عوامل برای بازشناسی پارامترهای مدل می‌باشد، مطالعه مفصلی جهت یافتن ساختار مناسب شبکه انجام شد. شبکه MLP در نظر گرفته شده، یک شبکه پیشخور است که در آن هر نرون در هر لایه به تمام نرون‌های لایه بعد متصل می‌باشد.^۱ پارامترهای مورد نظر برای طراحی یک شبکه، عبارتند از: تعیین نوع تابع محرک برای لایه‌های پنهان و خروجی که البته برای نرون‌های هر لایه یکسان فرض شده است، تعداد نرون‌ها در هر لایه پنهان و اثر اضافه کردن لایه پنهان دوم، انتخاب روش آموزشی مناسب، انتخاب روش مناسب نرمالایز کردن داده‌ها، انتخاب سطح خطای مناسب در فضای داده‌های نرمالایز و تقسیم بندی مناسب سری آموزشی به سری‌های آموزشی، اعتبارسنجی و آزمایشی، می‌باشد.

در یک شبکه MLP با یک لایه پنهان، خروجی محاسبه شده برای نرون k ام در لایه خروجی با در نظر گرفتن یک لایه پنهان برای شبکه به صورت زیر می‌باشد [ون در بین و جاتن، ۲۰۰۰].

$$o_k = f_k \left(\sum_{j=1}^m w_{kj} f_j \left(\sum_{i=1}^n w_{ji} x_i + w_{jb} \right) + w_{kb} \right) \quad (1-5)$$

o_k : خروجی نرون k ام

f_k : تابع محرک نرون k ام

w_{kj} : وزن‌های ارتباطی بین نرون‌های لایه پنهان و لایه خروجی

w_{kb} : جمله بایاس مربوط به نرون k ام در لایه خروجی

f_j : تابع محرک مربوط به نرون‌های لایه پنهان

m : تعداد نرون‌های لایه پنهان

1-fully connected

w_{ji} : وزن‌های ارتباطی بین نرون‌های لایه پنهان و ورودی

w_{jb} : جمله بایاس مربوط به نرون زام در لایه پنهان

x_i : بردار ورودی

n : تعداد ورودی‌ها

بدیهی است این معادله می‌تواند به آسانی برای نرون‌های دیگر لایه خروجی و همچنین برای لایه‌های پنهان بیشتر تعمیم یابد. بهتر است از تعداد کل متغیرهای داخلی شبکه n_{tot} که در تعیین رفتار شبکه عصبی مؤثر است، مطلع شویم.

خوشبختانه، محاسبه این تعداد با توجه به معادله (۵-۱) به آسانی میسر می‌باشد. برای شبکه MLP تعداد متغیرهای داخلی به صورت مجموع تعداد اتصالات وزنی و تعداد نرون‌ها در هر لایه که معادل با تعداد بایاس‌ها می‌باشد، محاسبه می‌شود.

چنانچه n_i ، تعداد ورودی‌ها، n_{hi} تعداد نرون‌ها در i امین لایه پنهان و n_o تعداد نرون‌های خروجی باشد، تعداد متغیرهای داخلی برای MLP با یک لایه پنهان به صورت

$$n_{tot} = (n_i + 1)n_{h1} + (n_{h1} + 1)n_o \quad (۲-۵)$$

و برای MLP با دو لایه پنهان به صورت

$$n_{tot} = (n_i + 1)n_{h1} + (n_{h1} + 1)n_{h2} + (n_{h2} + 1)n_o \quad (۳-۵)$$

می‌باشد.

چنانچه m تعداد نمونه‌های آموزشی باشد نسبت n_{tot}/m می‌تواند تا حدودی بیانگر عملکرد شبکه باشد. سه عامل می‌تواند دلیل عملکرد نامناسب یک شبکه باشد [رومئو، ۱۹۹۴] این عوامل عبارتند از: طراحی نامناسب شبکه، انتخاب الگوریتم آموزشی نامناسب و یا انتخاب سری آموزشی نامناسب. از مهمترین مسائل در طراحی شبکه، تعیین تعداد لایه‌های پنهان و نرون‌های هر لایه پنهان می‌باشد. هیچ قانونی برای تعیین تعداد نرون‌های لایه پنهان وجود ندارد. با این همه بررسی‌ها نشان می‌دهد که کران بالای تعداد نرون‌های مورد نیاز جهت تولید خروجی‌های مطلوب نمونه‌های آموزشی، در حدود m ، تعداد نمونه‌های آموزشی می‌باشد [هوانگ و هوانگ، ۱۹۹۱]. پس تعداد نرون‌ها در لایه پنهان هرگز نباید از تعداد نمونه‌های آموزشی فراتر رود و در ضمن تعداد

نمونه‌های آموزشی برای حصول به یک دقت خوب می‌بایست بزرگتر از تعداد وزن‌های داخلی باشد. در عمل حداکثر $m \approx 10n_{tot}$ و حداقل $m \approx 2 - 4n_{tot}$ پیشنهاد می‌شود [ون در بین و جاتن، ۲۰۰۰؛ پولتون، ۲۰۰۱]. بنابراین یک شبکه بزرگ نیاز به سری آموزشی بزرگی جهت ایجاد توازن بین تعداد وزن‌های ارتباطی و نمونه‌های آموزشی دارد [پولتون، ۲۰۰۱]. تعداد مناسب برای لایه‌های پنهان و نرون‌های لایه پنهان نیز توسط آزمون و خطا به دست می‌آید که معمولاً با یک تعداد کمی از نرون‌ها آغاز می‌شود. عامل دومی که باعث می‌شود شبکه به نتایج مطلوبی دست نیابد، به تله افتادن در کمینه‌های محلی می‌باشد. تابع خطا^۱ غالباً بسیار پیچیده می‌باشد [هاش^۲ و همکاران، ۱۹۹۲]. از این رو شبکه به جای حصول به کمینه سراسری ممکن است به آسانی در یک کمینه محلی به تله بیفتد. چاره کار، ساده می‌باشد چندین بار در جهت کمینه کردن خطا می‌بایست تلاش شود و هر بار نیز با استفاده از وزن‌دهی اولیه^۳ متفاوت (به طور تصادفی و یا غیر تصادفی) و یا از سایر الگوریتم‌های وارونه‌سازی مانند جستجوی سراسری^۳ استفاده شود. و سرانجام انتخاب سری‌های آموزشی نامناسب سبب عملکرد ضعیف شبکه می‌گردد. از معمولترین مشکلات آموزش شبکه می‌توان به مسأله برآزش بالا و سری آموزشی بد اشاره کرد. مورد دوم، شامل انتخاب الگوهای بد آموزشی می‌باشد. مثلاً چنانچه الگوها به کلاس اشتباه نسبت داده شوند یا سری آموزشی اجازه^۳ تعمیم خوبی به شبکه ندهد برای مثال فضای نمونه ممکن است ناکامل باشد [ون در بین و جاتن، ۲۰۰۰]. راه حل مشکل برآزش بالای شبکه، انتخاب مناسب تعداد نرون‌های پنهان، متناسب با تعداد نمونه‌های آموزشی می‌باشد [پولتون، ۲۰۰۱].

همان طور که قبلاً نیز ذکر شد یک شبکه جهت آموزش و دست‌یابی به قابلیت تعمیم‌دهی خوب به تعداد نمونه‌های آموزشی تقریباً ۱۰ برابر تعداد اتصالات وزنی نیاز دارد. به دلیل استفاده از شبکه‌های بزرگ در این تحقیق و عدم پاسخگویی سیستم‌های کامپیوتری در دسترس (Pentium 4-2.6GHz-512MB Ram) جهت آموزش شبکه، سری آموزشی انتخاب شده تنها دارای ۴۹۸ نمونه آموزشی می‌باشد. از این رو جهت تعمیم شبکه از سایر روش‌های تعمیم استفاده شده است.

1-misfit

2-Hush

3-global search

به منظور بررسی عملکرد شبکه از خطای RMS تعریف شده به صورت زیر استفاده شد. این خطا روی سری آزمایشی و به ازای تمام پارامترها تعریف شده است. [اسپیشاک و پوپوف، ۲۰۰۰].

$$RMS = \left[\frac{1}{N_{test} N_{par}} \sum_{n,j} err_{n,j}^2 \right]^{1/2} \quad (4-5)$$

$$err(n, j) = \frac{[target(n, j) - neural(n, j)]}{[\max(j) - \min(j)]}$$

$$j = 1, \dots, N_{par} \quad n = 1, \dots, N_{test}$$

که در رابطه فوق

$N_{par} (= 9)$: تعداد نرون‌های خروجی

N_{test} : تعداد نمونه‌های آزمایشی

$\max(j), \min(j)$ حداقل و حداکثر مقادیر زامین پارامتر در سری آموزشی

$neural(n, j)$ نتیجه پیش‌بینی شبکه برای زامین پارامتر در n امین نمونه آزمایشی $target(n, j)$ مقدار هدف برای زامین پارامتر در n امین نمونه آزمایشی می‌باشد. به منظور تخمین کیفیت وارونه‌سازی داده‌های مصنوعی توسط شبکه عصبی، برای هر پارامتر مجهول زام، میانگین خطای نسبی روی تمام نمونه‌های آزمایشی به صورت زیر محاسبه شد [اسپیشاک و پوپوف، ۲۰۰۰].

$$Err_j = \frac{1}{N_{test}} \sum_n \frac{|target_{n,j} - neural_{n,j}|}{target} \times 100\% \quad (5-5)$$

۵-۷- روش کار

جهت بررسی عملکرد و انتخاب برخی از پارامترهای شبکه مانند انتخاب نوع تابع محرک برای لایه پنهان و انتخاب روش آموزشی مناسب، نحوه ارائه ورودی‌ها به شبکه، نوع پیش‌پس پردازش مناسب و انتخاب سطح خطای مناسب به عنوان هدف^۱، ابتدا یک شبکه کوچکتر با انتخاب ۳۰ مدل به طور تصادفی از سری آموزشی ایجاد شد. جهت ارائه داده‌های ورودی به شبکه، داده‌های مصنوعی تولید شده به عنوان پاسخ هر مدل، به صورت بردار ستونی در ستون‌های یک ماتریس قرار گرفت بنابراین

1-goal

ماتریسی حاصل شد که تعداد ستون‌های آن معادل با تعداد مدل‌ها و تعداد سطرهای آن برابر با تعداد مؤلفه‌های اندازه‌گیری شده به ازای هر مدل بود.

متناظر با هر بردار ستونی قرار گرفته در ماتریس ورودی، یک بردار ستونی شامل پارامترهای مدل در ماتریس خروجی قرار گرفت، بنابراین پارامترهای مدل نیز تشکیل ماتریسی با تعداد سطرهای برابر با تعداد پارامترهای مدل (۹) و تعداد ستون‌های برابر با تعداد مدل‌های به کارگرفته برای آموزش شبکه، را داد.

برای ارائه تصویر کاملی از زمین، مؤلفه‌های مقاومت ویژه ظاهری و فاز امپدانس در هر دو مد که در ۹ ایستگاه به فواصل ۱۵۰۰ m و ۱۱ فرکانس اندازه‌گیری شده بود، به صورت چهار پارامتر ورودی به ترتیب ρ_{yx} , ρ_{yx} , ρ_{xy} , ρ_{xy} به صورت یک بردار ستونی قرار گرفت. هر پارامتر خود یک بردار ستونی متشکل از $9 \times 11 = 99$ مؤلفه می‌باشد. بنابراین هر ستون از ماتریس ورودی شامل $4 \times 99 = 396$ مؤلفه است.

پارامترهای خروجی نظیر هر مدل نیز به صورت ستون‌های یک ماتریس مرتب شد. بنابراین ابعاد ماتریس ورودی 396×30 و ماتریس خروجی 9×30 می‌باشد. اما اگر غیر از ورودی‌های مذکور، محل ایستگاه‌ها و مقادیر فرکانس یا پریود اندازه‌گیری نیز به ماتریس ورودی اضافه شود، آیا تأثیری در فرآیند آموزش دارد؟

این کار، بر روی شبکه کوچک آزمایش شد و مقادیر پریود و محل ایستگاه‌ها نیز به عنوان دو پارامتر و هر پارامتر شامل ۹۹ مؤلفه به همان ترتیب بیان شده به ماتریس ورودی اضافه شد. به این ترتیب هر بردار ورودی شامل ۶ پارامتر و هر پارامتر نیز شامل ۹۹ مؤلفه شد. این امر ابعاد ماتریس ورودی را به 594×30 افزایش داد. ولی دو پارامتر محل ایستگاه‌ها و مقادیر پریود در هر مدل تکرار می‌شدند، به بیان دیگر ۱۹۸ سطر از ماتریس ورودی عملاً بدون تغییر بودند، از این رو به دلیل ثابت ماندن این پارامترها در ماتریس ورودی، آن‌ها هیچ نقشی در فرآیند آموزش ندارند، علاوه بر این در این سطرها، فرآیند نرمالایز کردن داده‌ها نیز دچار مشکل می‌شد، بنابراین این دو پارامتر از مقادیر ورودی حذف شد.

۵-۷-۱- انتخاب روش نرمالایز کردن

شبکه مذکور یک بار بدون نرمالایز کردن داده‌های ورودی و خروجی، آموزش داده شد، که در این وضعیت خطای شبکه بسیار بالا و از رتبه 10^4 بود. از این رو آموزش داده‌ها در فضای نرمالایز با روش‌های مختلف آزمایش شد، انتخاب روش مناسب نرمالایز کردن داده‌ها بسیار مهم است به دلیل این که شبکه بر اساس داده‌های نرمالایز شده آموزش می‌بیند و در واقع به حداقل خطا در فضای داده‌های نرمالایز شده دست می‌یابد اما این داده‌ها باید دوباره به واحدهای دنیای واقعی^۱ برگردانده شوند و با این عمل خطا افزایش می‌یابد. از این رو روشی باید برای نرمالایز کردن داده‌ها، انتخاب گردد که در دنیای واقعی، خطای قابل قبولی داشته باشد. داده‌ها با استفاده از سه روش مختلف نرمالایز شدند.

۱. استفاده از الگوریتم ریاضی زیر که داده‌ها را به گونه‌ای نرمالایز می‌کند که انحراف معیار آن‌ها یک و میانگین صفر داشته باشند [راهنمای نرم افزار matlab. ۲۰۰۲].

$$p_n = \frac{(p - \text{mean } p)}{\text{std } p} \quad (5-6)$$

که در این رابطه p : ماتریس ورودی به ابعاد $R \times Q$

p_n : ماتریس ورودی‌های نرمالایز شده به ابعاد $R \times Q$

$\text{mean } p$: بردار $R \times 1$ شامل میانگین هر p و $\text{std } p$: بردار $R \times 1$ شامل انحراف معیار هر p می‌باشد.

این الگوریتم برای داده‌های ما چندان مناسب نبود.

۲. تقسیم مقادیر هر سطر به مقدار ماکزیمم آن سطر

استفاده از این روش نرمالایز کردن باعث می‌شد که شبکه به خطای بسیار پایینی در فضای نرمالایز دست‌یابد اما در برگرداندن داده‌ها به واحدهای دنیای واقعی خطا بسیار افزایش می‌یافت.

۳. استفاده از الگوریتم ریاضی زیر که داده‌ها را بین $[-1, 1]$ نرمالایز می‌کند [راهنمای نرم‌افزار

matlab. ۲۰۰۲].

$$p_n = \frac{2(p - \min p)}{(\max p - \min p)} - 1 \quad (5-7)$$

که در این رابطه p : ماتریس ورودی به ابعاد $R \times Q$

p : ماتریس ورودی‌های نرمالایز شده به ابعاد $R \times Q$

$minp$: بردار $R \times 1$ شامل مقدار کمینه به ازای هر p و $maxp$: بردار $R \times 1$ شامل مقدار بیشینه به ازای هر p می‌باشد این روش نرمالایز کردن نسبت به روش‌های قبل نتایج قابل قبولی را برای مطالعه حاضر ارائه داد.

۵-۷-۲- انتخاب نوع تابع فعالیت برای نرون‌های لایه پنهان و لایه خروجی

به دلیل این که وارونه‌سازی داده‌ها به عنوان یک مسأله تخمین تابع می‌باشد بهتر است از تابع محرک غیر خطی مانند سیگموئیدی برای نرون‌های لایه پنهان و از تابع خطی برای لایه خروجی استفاده شود. چنانچه از توابع سیگموئیدی برای نرون‌های لایه خروجی استفاده شود، خروجی‌های تولید شده توسط شبکه در محدوده معینی قرار می‌گیرند اما استفاده از تابع خطی در نرون‌های لایه خروجی این امکان را می‌دهد که خروجی‌های شبکه مقادیری خارج از یک گستره معین نیز داشته باشند [راهنمای نرم‌افزار matlab، ۲۰۰۲؛ پولتون، ۲۰۰۱].

در مورد انتخاب نوع تابع سیگموئیدی برای نرون‌های لایه پنهان هر دو تابع سیگموئیدی لگاریتمی (logsig) با الگوریتم ریاضی $\text{logsig}(x) = 1/(1+\exp(-x))$ و تانژانت هیپربولیک (tansig) با الگوریتم ریاضی $\text{tansig}(x) = 2/(1+\exp(-2x)) - 1$ در آموزش شبکه مورد آزمایش قرار گرفت که تابع logsig منجر به تولید نتایج بهتری شد.

به تدریج تعداد مدل‌های بیشتری به ماتریس ورودی و خروجی اضافه شد، با توجه به نتایج ارائه شده در جدول ۵-۱، ملاحظه می‌شود، با افزایش تعداد نمونه‌های آموزشی مناسب به شبکه، دقت تخمین پارامترهای مدل افزایش و شبکه به خطای RMS کمتری دست می‌یابد ولی همانطور که قبلاً نیز اشاره شد چنانچه نمونه‌های آموزشی نامناسبی به شبکه ارائه شود خطا نه تنها کاهش نمی‌یابد بلکه افزایش نیز می‌یابد. لازم به ذکر است در جدول ۵-۱ تعداد کل نمونه‌ها ارائه شده است که شامل نمونه‌های آموزشی، اعتبارسنجی و آزمایشی می‌باشد. البته تقسیم بندی نمونه‌ها در مورد تمام شبکه‌های ذکر شده یکسان می‌باشد، در این تقسیم بندی ۲۰٪ از کل نمونه‌ها در سری اعتبارسنجی، ۱۰٪ در سری آزمایشی و بقیه نمونه‌ها در سری آموزشی مورد استفاده قرار گرفته شده است.

جدول ۵-۱- تغییرات خطای RMS با افزایش تعداد نمونه‌ها

تعداد کل مجموعه نمونه‌های ارائه شده به شبکه	خطای RMS
124	0.29
155	0.20
185	0.26
498	0.156

نهایتاً شبکه نسبتاً بزرگی با ماتریس ورودی به ابعاد 396×498 و ماتریس خروجی به ابعاد 9×498 تشکیل شد. سطح خطای هدف در فضای نرمالایز نیز 0.001 در نظر گرفته شد. از آن جایی که طراحی یک شبکه و انتخاب مقادیر مناسب پارامترهای شبکه غالباً از طریق آزمون و خطا حاصل می‌شود، طراحی ساختار مناسب برای یک شبکه بزرگ با توجه به امکانات کامپیوتری در دسترس بسیار وقت‌گیر بود.

۵-۷-۳- انتخاب روش آموزشی مناسب

الگوریتم آموزشی پس انتشار خطا به صورت یک روش کاهش شیب^۱ عمل می‌کند. ضعف عمده این الگوریتم، سرعت کم همگرایی آن و نیز توقف یادگیری شبکه به دلیل به دام افتادن پارامترهای شبکه در نقاط کمینه محلی می‌باشد.

در روش‌های کاهش شیب، مقدار نرخ یادگیری در مدت آموزش ثابت نگه‌داشته می‌شود. به بیان دیگر در الگوریتم پس انتشار خطا وزن‌ها همواره با یک ضریب ثابت (نرخ یادگیری) از مشتق جزئی خطا نسبت به وزن‌ها، تصحیح می‌شوند. بنابراین چنانچه سطح خطا کاملاً صاف باشد، مشتق کوچک می‌شود و در نتیجه تغییرات وزن‌ها نیز کوچک می‌گردند، تغییرات وزنی کوچک، زمان همگرایی را افزایش می‌دهد چنانچه سطح خطا شدیداً پرپیچ و خم باشد مشتق می‌تواند بزرگ شود که منجر به تغییرات وزنی بزرگ می‌شود. تغییرات وزنی بزرگ احتمال گم کردن (از دست دادن) کمینه سراسری را افزایش می‌دهد. بنابراین عملکرد الگوریتم مذکور به انتخاب مناسب نرخ یادگیری بسیار حساس می‌باشد. انتخاب مناسب نرخ یادگیری قبل از آموزش، عملی نیست چنانچه ما اجازه

¹-gradient descent

تغییر به نرخ یادگیری در خلال فرآیند آموزش بدهیم، عملکرد الگوریتم کاهش می‌تواند توسعه یابد [منہاج، ۱۳۷۹؛ پولتون و همکاران، ۱۹۹۲].

۵-۷-۳- الف - روش آموزشی لوبنبرگ - مارکوآرت^۱

روش‌های آموزشی مختلفی وجود دارد که در آن‌ها نرخ یادگیری در طول فرآیند آموزش تغییر می‌کند چنانچه روش پس انتشار خطا به صورت زیر نوشته شود.

$$x_{k+1} = x_k - \alpha_k g_k \quad (۸-۵)$$

که در آن:

x_k : بردار شامل وزن‌ها و بایاس‌ها در تکرار k ام

g_k : گرادیان بردار خطا نسبت به وزن‌ها و بایاس‌ها و α_k : نرخ یادگیری می‌باشد

اگر $g = J^T e$ نوشته شود، که در آن e بردار خطاهای شبکه می‌باشد و J ماتریس ژاکوبین شامل مشتقات اول خطاهای شبکه نسبت به وزن‌ها و بایاس‌ها می‌باشد، در الگوریتم لوبنبرگ-مارکوآرت نرخ یادگیری به صورت تقریبی از ماتریس هسیان محاسبه می‌گردد. محاسبه ماتریس هسیان به صورت $H = J^T J$ تقریب می‌شود از این رو در این الگوریتم محاسبه ماتریس هسیان از محاسبات و پیچیدگی کمتری برخوردار است. اصلاح وزن‌ها و بایاس‌ها در این الگوریتم به صورت زیر می‌باشد:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \quad (۹-۵)$$

که در رابطه فوق ماتریس واحد است [منہاج، ۱۳۷۹؛ راهنمای نرم افزار مطلب، ۲۰۰۲].

این الگوریتم تنها از آموزش گروهی می‌تواند استفاده کند و هیچ تضمینی ندارد که با آموزش انگو به انگو همگرا شود، البته محاسبه ماتریس هسیان حافظه زیادی را اشغال می‌کند. اما از خصوصیات مهم این الگوریتم سرعت همگرایی فوق العاده آن است و از سریعترین الگوریتم‌های آموزشی می‌باشد [منہاج، ۱۳۷۹]. در این مطالعه، سایر الگوریتم‌های آموزشی نیز آزمایش شد که با توجه به قابلیت‌های الگوریتم لوبنبرگ-مارکوآرت، از این الگوریتم استفاده شد.

۵-۷-۴- تعمیم شبکه

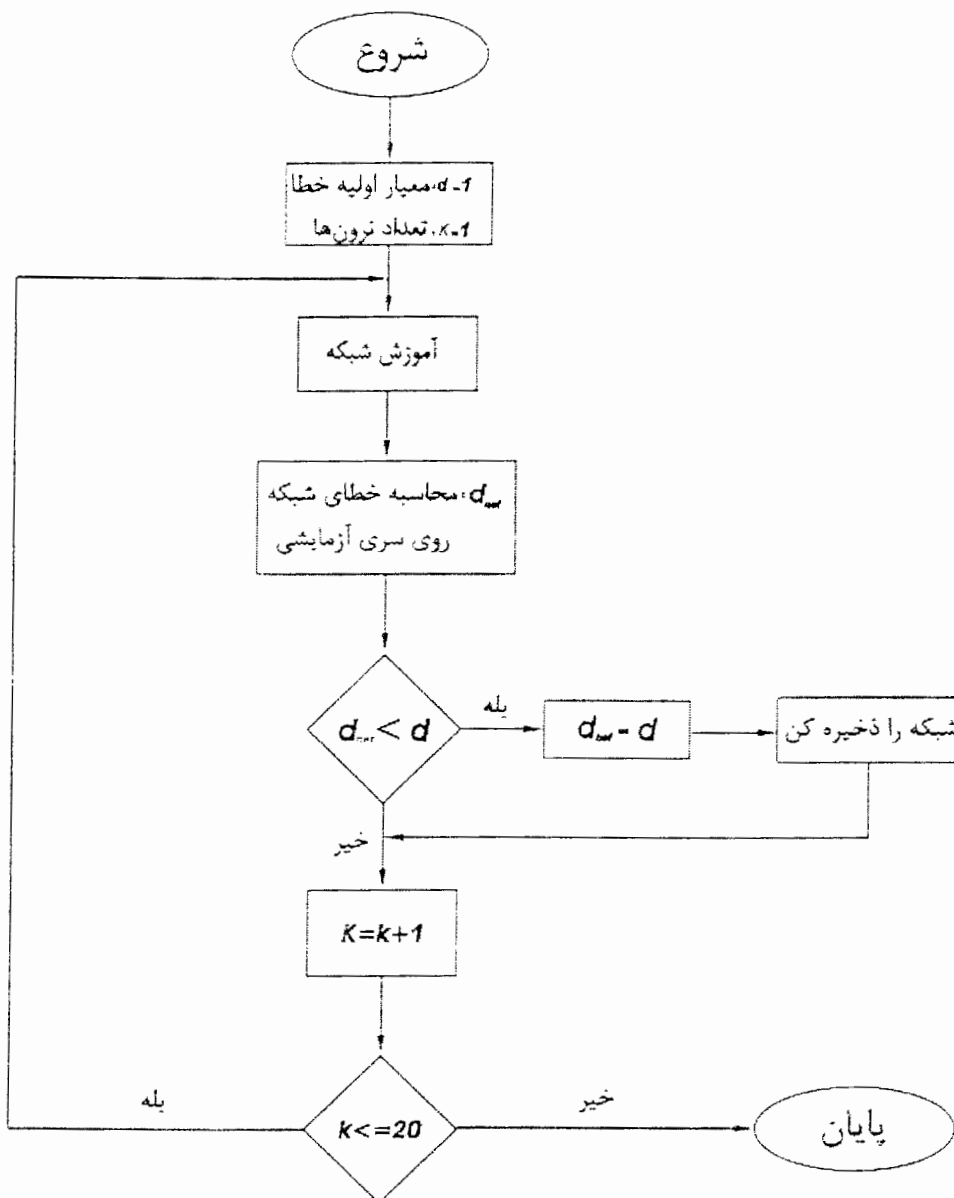
جهت انتخاب روش مناسب برای تعمیم دهی شبکه دو روش مورد آزمایش قرار گرفت در اولین روش سری آموزشی به دو زیر شاخه تقسیم شد. در این تقسیم بندی ۱۰ درصد از نمونه‌های آموزشی به طور تصادفی جهت تشکیل سری آزمایشی انتخاب شدند و بقیه نمونه‌ها نیز برای آموزش مورد استفاده قرار گرفتند. ساختار مناسب انتخاب شده برای این شبکه به صورت ۹-۱۲-۳۹۶ می‌باشد که براین اساس این شبکه سه لایه‌ای دارای ۳۹۶ نرون در لایه ورودی، ۱۲ نرون در لایه پنهان و ۹ نرون در لایه خروجی می‌باشد. جهت دستیافتن به تعداد مناسب نرون‌ها در لایه پنهان با توجه به محدودیت سیستم‌های کامپیوتری در دسترس و بزرگ بودن سری آموزشی (۴۹۸ مدل) مدت زمان ۷۲ ساعت صرف شد، این شبکه پس از آموزش به خطای $RMS=0/234$ دست یافت. با توجه به بزرگ بودن شبکه و با در نظر گرفتن رابطه (۵-۲)، این شبکه برای تعمیم دهی خوب نیاز به تعداد زیادی نمونه آموزشی دارد.

روش دوم استفاده از روش توقف سریع می‌باشد، که در این روش جهت تعمیم شبکه از تقسیم سری آموزشی به سه دسته آموزشی، اعتبار سنجی و آزمایشی استفاده شد در این تقسیم بندی ۲۰٪ از کل نمونه‌های آموزشی به طور تصادفی به عنوان سری اعتبار سنجی، ۱۰٪ به عنوان سری آزمایشی و ۷۰٪ نیز به عنوان سری آموزشی انتخاب شد. اشکال این روش این است که با در نظر گرفتن بخشی از سری آموزشی برای سری اعتبارسنجی از تعداد نمونه‌های آموزشی کم می‌شود اما به دلیل تغییر شرط توقف، مدت زمان صرف شده برای آموزش این شبکه‌ها کاهش می‌یابد. این شبکه پس از آموزش به $RMS=0/156$ دست یافت. به دلیل محدودیت زمانی و عدم دسترسی به کامپیوترهای سریع روش منظم سازی داده‌ها آزمایش نشد. با توجه به خطاهای RMS حاصل شده و مدت زمان صرف شده برای آموزش، جهت تعمیم شبکه، روش توقف سریع انتخاب شد.

۵-۷-۵ - تعیین تعداد لایه های پنهان و تعداد نرون‌های لایه پنهان

ابتدا یک شبکه سه لایه‌ای، شامل یک لایه ورودی، یک لایه پنهان و یک لایه خروجی در نظر گرفته شد. برای تعیین تعداد نرون‌های لایه پنهان، برنامه آموزش شبکه به ازای تعداد نرون‌های پنهان

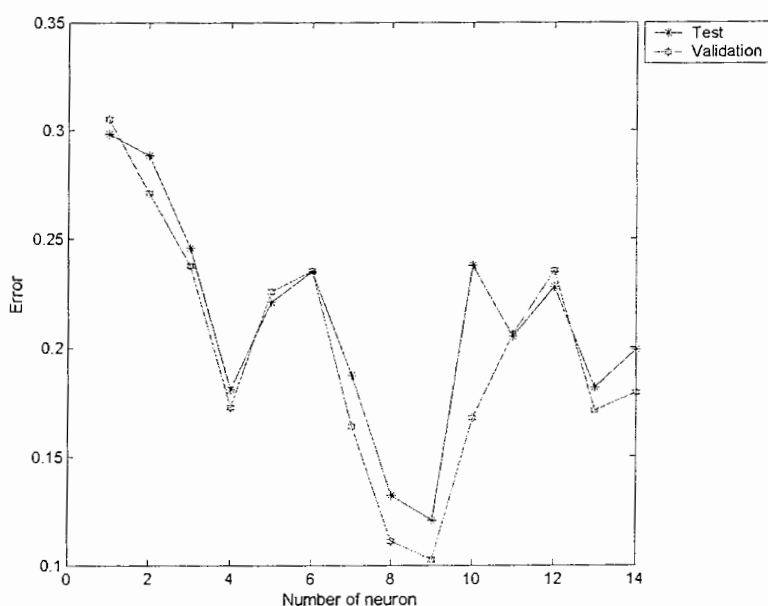
مختلف مطابق فلوجارت زیر در یک حلقه تکرار قرار گرفت. با توجه به بزرگ بودن شبکه و امکانات کامپیوتری موجود حداکثر تعداد نرون‌های پنهان برای این شبکه حدود ۲۰ نرون در نظر گرفته شد.



شکل ۵-۶- فلوجارت تعیین تعداد نرون‌های پنهان

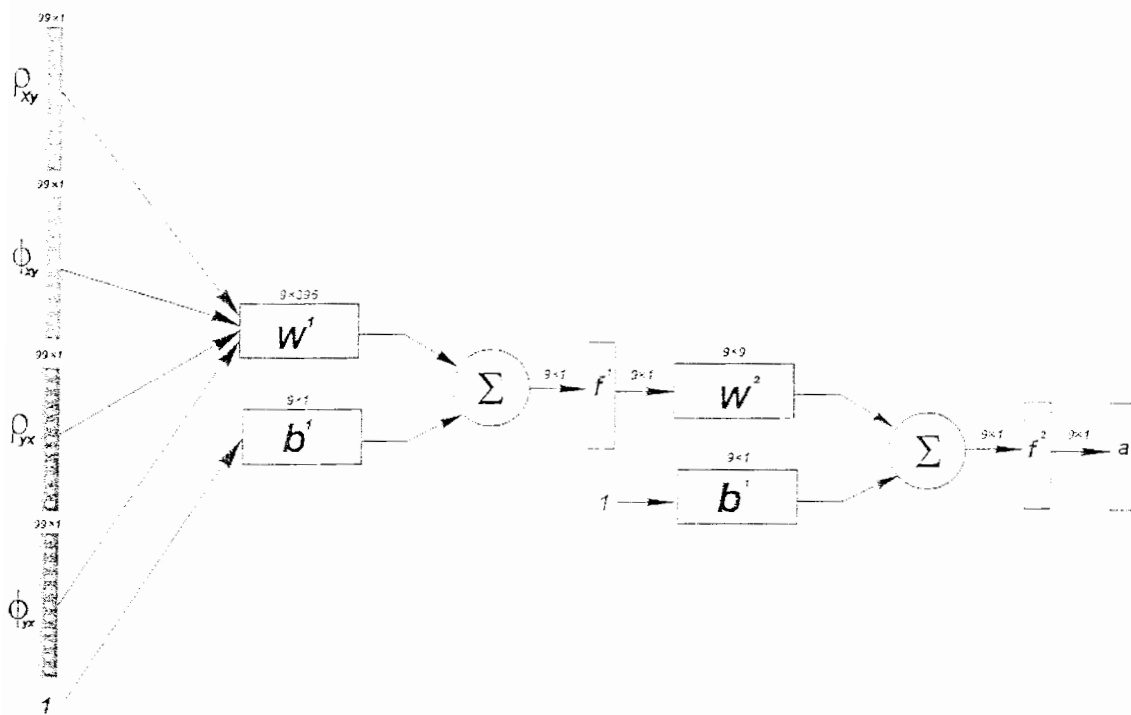
این الگوریتم با قرار دادن یک نرون در لایه پنهان شبکه و انتخاب $d=1$ به عنوان معیار اولیه خطا شروع می‌شود. وزن‌ها و بایاس‌های اولیه شبکه صفر انتخاب می‌شود. پس از آموزش شبکه، مقدار خطا برای سری‌های آموزشی، اعتبار سنجی و آزمایشی محاسبه می‌شود. مقدار خطای سری‌های اعتبار سنجی و آزمایشی برای ما مهم است چون این نمونه‌ها در فرآیند آموزش شبکه شرکت

نمی‌کنند. خطای شبکه روی سری آزمایشی (d_{net}) با معیار اولیه خطا مقایسه می‌شود چنانچه $d_{net} < d$ باشد d_{net} جایگزین d می‌شود و وزن‌ها و بایاس‌های به دست آمده پس از آموزش شبکه ذخیره می‌شوند. سپس به تعداد نرون‌های پنهان یکی اضافه می‌شود. تعداد جدید نرون‌های پنهان با حداکثر این تعداد مقایسه می‌شود. چنانچه تعداد انتخاب شده نرون‌ها کمتر از حداکثر تعداد تعیین شده باشد، مرحله آموزش شبکه و محاسبه خطا تکرار می‌شود. چنانچه $d_{net} > d$ بدون قرار دادن d_{net} به جای d و ذخیره‌سازی شبکه، به تعداد نرون‌های پنهان یکی اضافه می‌شود و به همان ترتیب بیان شده در بالا، روند ادامه می‌یابد. لازم به ذکر است، در هر تکرار و پیش از آموزش شبکه، مقادیر اولیه وزن‌ها و بایاس‌ها به مقدار صفر برگردانده می‌شوند. هر چند حداکثر تعداد تعیین شده برای نرون‌های پنهان در این شبکه ۲۰ نرون در نظر گرفته شد ولی سیستم کامپیوتری موجود (Pentium 4-2.6GHz-512MB Ram) برای تعداد ۱۴ نرون پنهان، پیغام کمبود حافظه را ثبت کرد. باتوجه به شکل ۵-۷ و مقادیر ثبت شده خطا برای سری آزمایشی و اعتبارسنجی انتخاب تعداد ۹ نرون برای اولین لایه پنهان، مناسب به نظر می‌آید. شبکه مذکور که شامل الگوی ورودی پارامترهای مقاومت ویژه ظاهری و فاز امپدانس در هر دو مد می‌باشد، برای سهولت تحت عنوان شبکه MT نامگذاری می‌گردد.



شکل ۵-۷- نمودار تعیین تعداد نرون‌های پنهان شبکه MT

برای اضافه کردن دومین لایه پنهان نیز برنامه آموزش شبکه به ازای تعداد مختلف نرون‌های پنهان در لایه دوم، در یک حلقه تکرار قرار گرفت. در این فرآیند تعداد نرون‌های اولین لایه پنهان ۹ نرون و ثابت در نظر گرفته شد. این بار حداقل خطای به دست آمده برای شبکه ۳ لایه به عنوان معیاری جهت تعیین بنبود یا عدم بنبود دقت شبکه با اضافه کردن یک لایه پنهان دیگر در نظر گرفته شد. اما تنها با اضافه کردن ۱ نرون در دومین لایه پنهان، سیستم کامپیوتری در دسترس پس از گذشت ۲۰ ساعت هنوز پاسخی مثبت نکرده و آموزش شبکه متوقف نشده بود. از این رو همان ساختار ۹-۹-۳۹۶ برای شبکه MT در نظر گرفته شد. در شکل ۵-۸ ساختار شبکه MT نشان داده می‌شود.

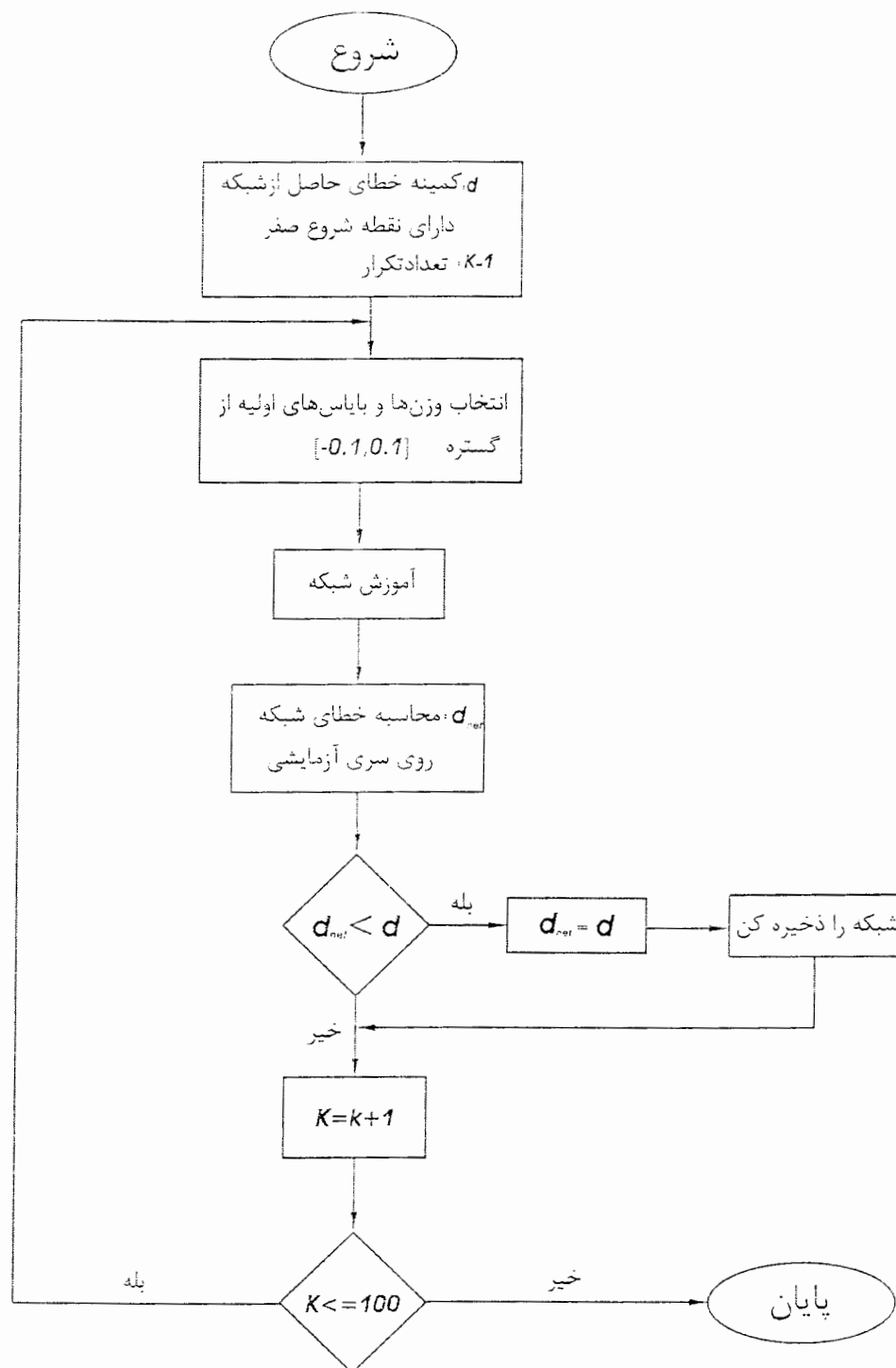


شکل ۵-۸- ساختار شبکه MT به صورت خلاصه شده

۵-۶- مقدار دهی اولیه برای وزن‌ها و بایاس‌ها

مقدار دهی اولیه و مناسب وزن‌ها و بایاس‌ها به عنوان یک نقطه شروع خوب می‌تواند در افزایش دقت و سرعت همگرایی شبکه مؤثر باشد در غیر این صورت ممکن است فرآیند آموزش بسیار آهسته صورت گیرد و خطر به دام افتادن در کمینه‌های محلی نیز افزایش یابد.

معمول‌ترین روش مقدار دهی اولیه به وزن‌ها و بایاس‌ها، انتخاب آن‌ها به صورت تصادفی مثلاً $w_i e^{-r, r}$ می‌باشد [ون در بین و جاتن، ۲۰۰۰]. مقادیر مختلفی برای r مورد آزمایش قرار گرفت که مقدار مناسب $r = 0.1$ انتخاب شد. برای به دست آوردن نقطه شروع مناسب برای وزن‌ها و بایاس‌ها نیز، آموزش شبکه در یک حلقه تکرار قرار گرفت و این بار حداقل خطای به دست آمده روی سری آزمایشی حاصل از ساختار تعیین شده شبکه با نقطه شروع صفر به عنوان معیاری برای مقایسه بهبود یا عدم بهبود دقت شبکه با مقدار دهی اولیه وزن‌ها و بایاس‌ها در گستره مورد نظر، در نظر گرفته شد. فلوچارت ارائه شده در شکل ۵-۹ چگونگی حصول نقطه شروع مناسب را نشان می‌دهد. بر طبق این الگوریتم، کمینه خطای به دست آمده از آموزش شبکه با مقدار اولیه وزن‌ها و بایاس‌های صفر به عنوان معیار اولیه خطا در نظر گرفته شد. در ضمن مقدار $k=1$ به عنوان شماره مرحله تکرار تعیین شد. که حداکثر تعداد تکرار در این الگوریتم ۱۰۰ بار می‌باشد. ماتریس وزنی با ابعاد 9×396 بین لایه ورودی و پنهان، بردار بایاس 9×1 برای نرون‌های پنهان، ماتریس وزنی 9×9 بین لایه پنهان و خروجی و بردار 9×1 نیز برای نرون‌های خروجی با انتخاب اعداد تصادفی از گستره $[-0.1, 0.1]$ ، به عنوان وزن‌ها و بایاس‌های اولیه در شبکه MT تشکیل شدند. سپس شبکه با نقطه شروع انتخاب شده، آموزش می‌بیند. مقدار خطای شبکه برای سری‌های آموزشی، اعتبارسنجی و آزمایشی محاسبه می‌شود. خطای محاسبه شده برای سری آزمایشی (d_{net}) با خطای در نظر گرفته شده به عنوان معیار اولیه (کمینه مقدار خطای حاصل با نقطه شروع صفر: d) مقایسه می‌شود. چنانچه $d_{net} < d$ باشد جایگزین d می‌شود و وزن‌ها و بایاس‌های به دست آمده پس از آموزش شبکه ذخیره می‌شوند. سپس به تعداد مرحله‌های تکرار یکی اضافه می‌شود. تعداد جدید مرحله‌های تکرار با حداکثر این تعداد مقایسه می‌شود. چنانچه تعداد انتخاب شده مرحله‌های تکرار کمتر از حداکثر تعداد تعیین شده باشد، مرحله آموزش شبکه و محاسبه خطا تکرار می‌شود. چنانچه $d_{net} > d$ بدون قرار دادن d_{net} به جای d و ذخیره‌سازی شبکه، به تعداد مرحله‌های تکرار یکی اضافه می‌شود و به همان ترتیب بیان شده در بالا، روند ادامه می‌یابد. لازم به ذکر است، در هر تکرار و پیش از آموزش شبکه، مقادیر اولیه وزن‌ها و بایاس‌ها به همان ترتیب بیان شده به طور تصادفی از گستره $[-0.1, 0.1]$ انتخاب می‌شوند.

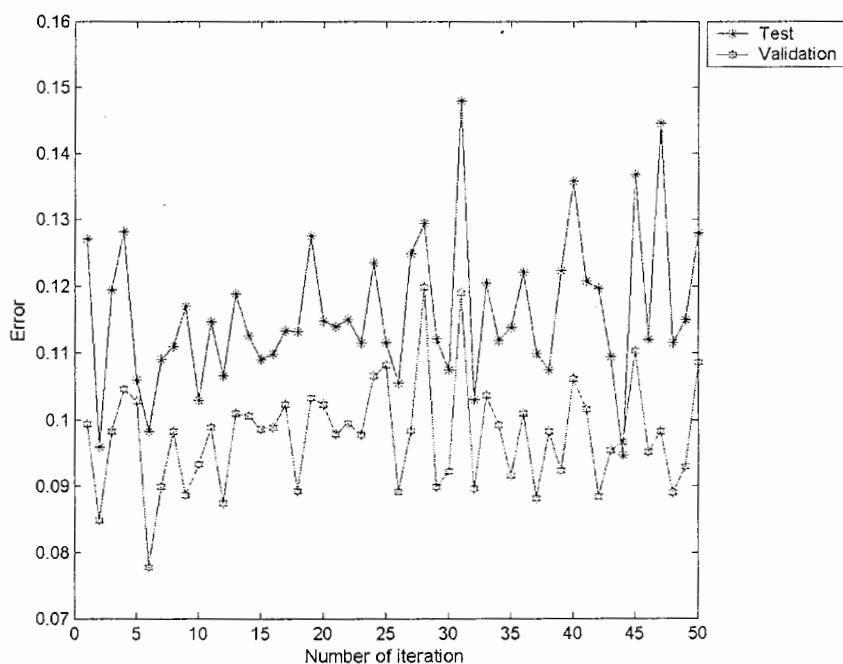


شکل ۵-۹- فلوجارت تعیین نقطه شروع مناسب

در شکل ۵-۱۰ نمودار تعیین نقطه شروع برای ۵۰ بار تکرار در شبکه MT، ارائه می‌شود. همان‌طور که در این شکل مشاهده می‌شود نوسانات خطا بین ۰.۰۹ و ۰.۱۵ روی سری آزمایشی حاکی حساسیت شبکه به انتخاب مقادیر اولیه وزن‌ها و بایاس‌ها می‌باشد. ملاحظه می‌شود در ۴۴ امین تکرار

خطای سری آزمایشی به حداقل می‌رسد. بنابراین وزن‌ها و بایاس‌های به دست آمده پس از آموزش شبکه با این نقطه شروع ذخیره می‌گردد.

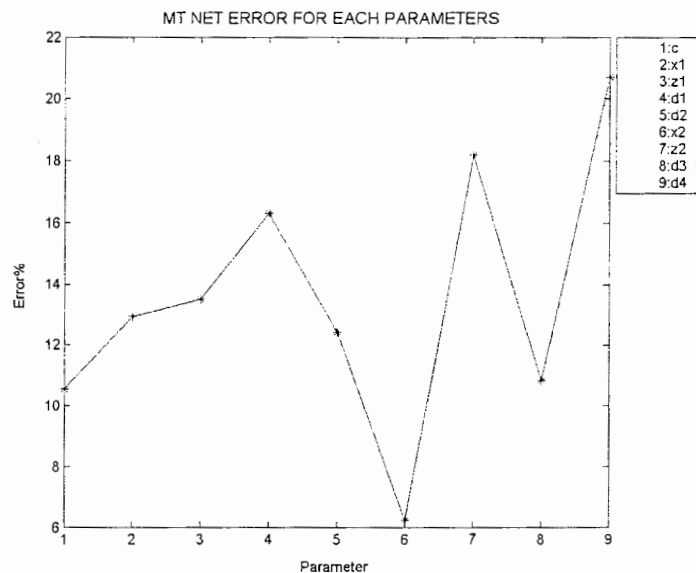
لازم به ذکر است مقدار خطای در نظر گرفته شده به عنوان معیار انتخاب در هر سه مورد تعیین تعداد لایه‌های پنهان، تعداد نرون‌های پنهان و انتخاب نقطه شروع وزن‌ها و بایاس‌ها، مقدار خطای سری آزمایشی می‌باشد. به این ترتیب ساختار مناسب برای شبکه MT طراحی می‌شود.



شکل ۵-۱۰- نمودار تعیین نقطه شروع مناسب برای وزن‌ها و بایاس‌ها در شبکه MT

خطای RMS محاسبه شده با استفاده از معادله (۴-۵) برای شبکه MT برابر با ۰/۱۵۶ به دست آمد. شکل ۵-۱۱ میانگین خطای نسبی محاسبه شده از طریق رابطه (۵-۵) را به ازای هر پارامتر نشان می‌دهد. مقادیر درج شده در ستون کنار شکل مقادیر عددی خطا به ازای هر پارامتر را نشان می‌دهد.

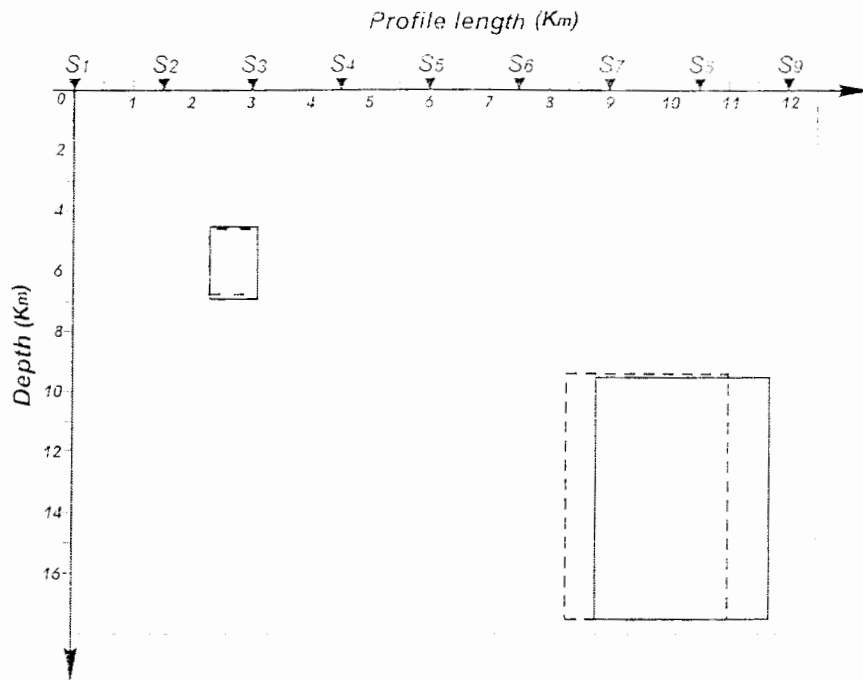
پارامتر	درصد خطا
c	10.546
x_1	12.917
z_1	13.507
d_1	16.307
d_2	12.45
x_2	6.2278
z_2	18.2
d_3	10.859
d_4	20.712



شکل ۵-۱۱- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه MT به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ

همان طور که مشاهده می‌شود کمترین خطای به دست آمده مربوط به پارامتر ششم (x_2) و بیشترین خطا مربوط به پارامتر نهم (d_4) می‌باشد. علیرغم پوشش یکنواخت فضای مربوط به تمام پارامترها ملاحظه می‌شود که میزان خطا در مورد پارامترهای مختلف، متفاوت است. علت این امر را باید در نحوه یادگیری شبکه جستجو کرد. شبکه با استخراج مشخصه‌های بارز الگوهای ورودی سعی می‌کند با استفاده از الگوریتم آموزشی تعیین شده، رابطه بین این مشخصه‌ها را با پارامترهای خروجی به دست آورد. هر یک از پارامترهای خروجی تابعی از این مشخصه‌ها می‌باشند. تقریب یک تابع برای تمام پارامترها که به وسیله آن بتوان در یک زمان ۹ پارامتر را تخمین زد موجب خواهد شد تا تابع تقریب شده توسط شبکه برای تخمین یکی از پارامترها مناسب‌تر از پارامتر دیگر باشد و یا به بیان دیگر میزان حساسیت شبکه به یک پارامتر بیشتر از پارامترهای دیگر باشد.

جهت بررسی بهتر عملکرد شبکه، پارامترهای مدل واقعی و پارامترهای مدل تخمین زده شده توسط شبکه، برای یکی از نمونه‌های آزمایشی در شکل ۵-۱۲ رسم شده است، مدل نشان داده شده با خط توپر نمایانگر مدل واقعی و مدل نشان داده شده با خط چین بیانگر پاسخ شبکه می‌باشد. مقادیر عددی پارامترهای واقعی و تخمین زده شده نیز در جدول ۵-۲ ارائه شده است. لازم به ذکر است واحد پارامترهای هندسی ارائه شده در تمامی جداول، متر می‌باشد.



شکل ۵-۱۲- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه MT (خط چین)

جدول ۲-۵- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه MT برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	50	53.95
x_1	2700	2691.8
z_1	5700	5667.8
d_1	400	407.96
d_2	1200	1074
x_2	10250	9659.5
z_2	13500	13446
d_3	1450	1348.6
d_4	4000	4062

با توجه به نتایج مندرج در جدول ۲-۵ و همچنین با توجه به شکل ۵-۱۲، ملاحظه می‌شود این شبکه از آموزش و تعمیم خوبی برخوردار بوده است به طوری که با ارائه نمونه‌ای که شبکه آن را در طی فرآیند آموزش ندیده است (سری آزمایشی) غیر از محل مرزهای جانبی در آنومالی سمت راست، بقیه پارامترها از دقت قابل قبولی برخوردار می‌باشد. در کل مقدار میانگین خطای نسبی روی

تمام پارامترها ۳/۹۹۸ درصد می‌باشد که این امر حاکی از دقت بالای شبکه MT در تخمین پارامترهای مدل می‌باشد.

۵-۸- اثر انتخاب نوع داده‌های ورودی

نتایج حاصل از آموزش شبکه با توجه به نوع ورودی‌ها مورد بررسی قرار گرفت. براین اساس چهار شبکه دیگر نیز به همان ترتیب بیان شده با ورودی‌های مختلف طراحی و آموزش داده شد این شبکه‌ها عبارتند از:

(۱) شبکه RES با مقادیر مقاومت ویژه ظاهری ρ_{yx}^a و ρ_{xy}^a به عنوان ورودی

(۲) شبکه PHASE با مقادیر فاز φ_{yx} و φ_{xy} به عنوان ورودی

(۳) شبکه TE با مقادیر مؤلفه‌های مقاومت ویژه ظاهری و فاز، ρ_{xy}^a و φ_{xy} (مد TE) به عنوان

ورودی

(۴) شبکه TM با مقادیر مؤلفه‌های مقاومت ویژه ظاهری و فاز ρ_{yx}^a و φ_{yx} (مد TM) به عنوان

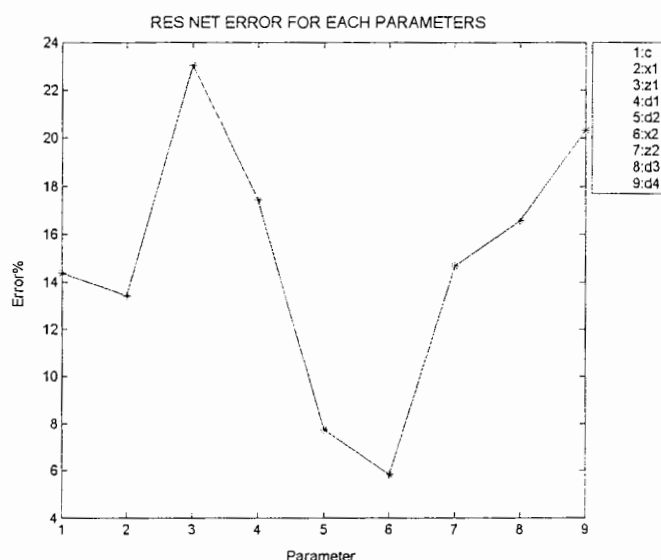
ورودی

۵-۸-۱- شبکه RES

مقادیر ورودی شبکه RES به همان ترتیب بیان شده به صورت بردارهای ستونی شامل (2×99) مقاومت ویژه ظاهری ρ_{yx}^a و ρ_{xy}^a برای مدهای TE و TM بوده است. این بار شبکه کوچکتری با ماتریس ورودی به ابعاد 198×498 و ماتریس خروجی به ابعاد 9×498 ایجاد شد به دلیل کاهش ابعاد، سرعت آموزش شبکه افزایش یافت. ساختار مناسب طراحی شده برای این شبکه به صورت ۹-۱۷-۱۳-۱۹۸ می‌باشد.

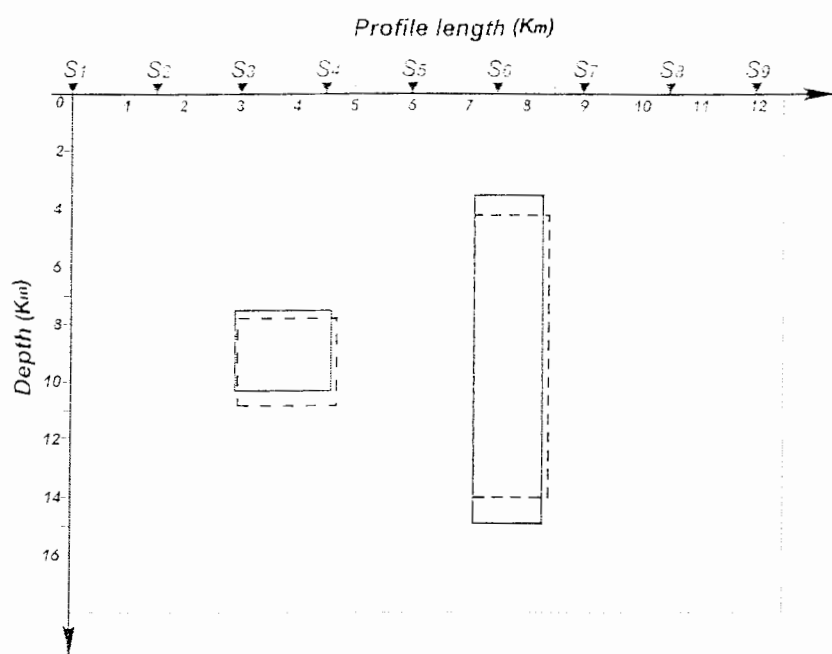
مقدار خطای RMS به دست آمده در این شبکه برابر با ۰/۱۵ می‌باشد. نمودار تغییرات میانگین خطای نسبی به دست آمده روی سری آزمایشی به ازای هر پارامتر در شکل ۵-۱۳ ارائه شده است همچنین مقادیر عددی میانگین خطای نسبی نیز در ستون کنار شکل درج شده است.

پارامتر	درصد خطا
c	14.362
x_1	13.406
z_1	23.035
d_1	17.419
d_2	7.7504
x_2	5.8515
z_2	14.68
d_3	16.562
d_4	20.341



شکل ۵-۱۳- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه RES به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ

با توجه به نمودار و نتایج مندرج در جدول فوق، مشاهده می‌شود کمترین خطا مربوط به پارامتر ششم (x_2) و بیشترین خطا در این شبکه مربوط به پارامتر سوم (z_1) می‌باشد. به همان دلایلی که در مورد شبکه MT نیز ذکر شد، در این جا تابع تقریب شده توسط شبکه RES برای تخمین پارامترهای پنجم و ششم نسبت به پارامترهای دیگر مناسب‌تر می‌باشد. برای ارزیابی بهتر نتایج، مدل واقعی و تخمین زده شده توسط این شبکه برای یکی از نمونه‌های آزمایشی که در فرآیند آموزش شرکت نداشته‌اند، در شکل ۵-۱۴ نشان داده می‌شود، همچنین مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES در جدول ۵-۳ ارائه می‌شود. همان طور که در شکل مشاهده می‌شود و با توجه به نتایج مندرج در جدول ۵-۳ شبکه در تخمین پارامترهای مرزهای بالایی، پایینی و سمت راست هر دو آنومالی به میزان کمی دچار خطا می‌شود ولی در مورد تخمین محل آنومالی‌ها نسبتاً دقیق عمل می‌کند. میانگین خطای نسبی روی تمام پارامترها ۴/۹۳۶ درصد می‌باشد که بیانگر دقت خوب این شبکه در تخمین پارامترهای مدل می‌باشد.



شکل ۵-۱۴- منایسه مدل واقعی (خط پر) و مدل تخمینی توسط شبکه RES (خط چین)

جدول ۵-۳- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES برای یکی از نمونه‌های آزمایشی

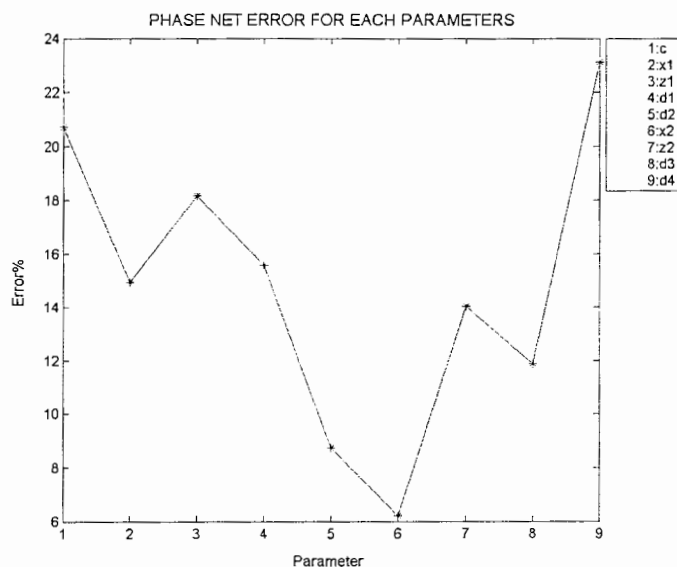
پارامترهای مدل	مدل واقعی	مدل تخمین زده شده
c	70	71.277
X_1	3750	3824.6
Z_1	8900	9286.6
d_1	850	873.19
d_2	1400	1523.5
X_2	7700	7761.9
Z_2	9200	9099.4
d_3	600	653.19
d_4	5700	4904.5

۵ - ۸ - ۲- شبکه PHASE

این شبکه شامل ورودی فاز هر دو مد (φ_{yx} و φ_{xy}) می‌باشد که هر یک از آنها متشکل از ۹۹ مؤلفه است. بنابراین ماتریس ورودی دارای ابعاد ۱۹۸×۴۹۸ و ماتریس خروجی نیز ۹×۴۹۸ می‌باشد. ساختار طراحی شده برای شبکه ۹-۱۲-۱۹۸ می‌باشد. مقدار خطای RMS به دست آمده در این شبکه برابر با ۰/۱۶ می‌باشد. میانگین خطای نسبی به دست آمده روی سری آزمایشی به ازای هر پارامتر در شکل ۵-۱۵ نشان داده شد و همچنین مقادیر عددی

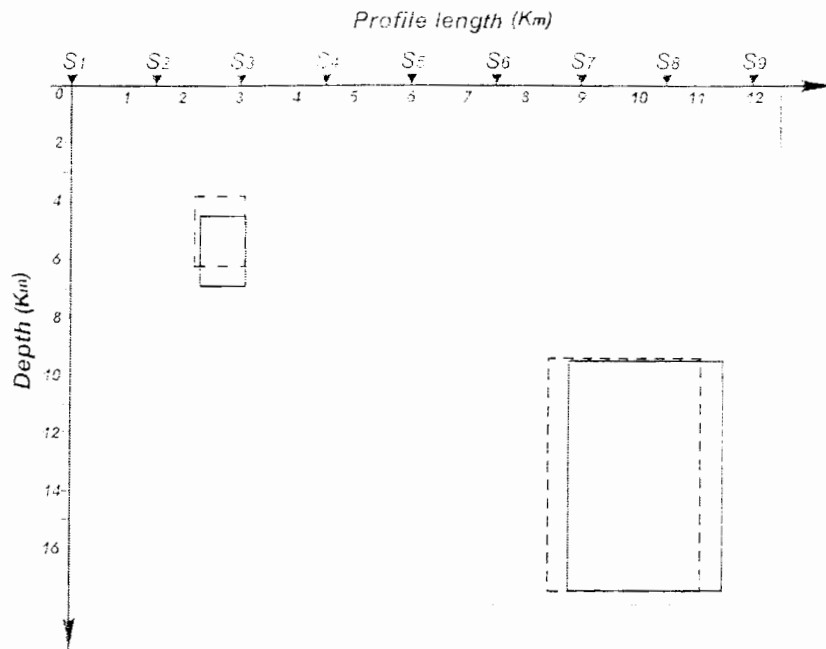
آن‌ها نیز در ستون کنار شکل ارائه شده است. با توجه به نتایج به دست آمده ملاحظه می‌شود کمینه خطای شبکه PHASE، همانند شبکه RES مربوط به پارامترهای پنجم و ششم می‌شود و بیشینه خطا مربوط به پارامتر نهم می‌شود.

پارامتر	درصد خطا
c	20.723
x ₁	14.939
z ₁	18.18
d ₁	15.572
d ₂	8.7351
x ₂	6.2194
z ₂	14.022
d ₃	11.881
d ₄	23.119



شکل ۵-۱۵- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه PHASE به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ

در ضمن برای مقایسه بهتر، مدل واقعی و مدل تخمینی توسط شبکه، به ازای یکی از نمونه‌های آزمایشی در شکل ۵-۱۶ نشان داده می‌شود، همچنین مقادیر عددی پارامترهای واقعی مدل به همراه پارامترهای تخمین زده شده توسط شبکه در جدول ۵-۴ ارائه شده است. با توجه به شکل و نتایج مندرج در جدول ۵-۴ مشاهده می‌شود که این شبکه در مورد تخمین پارامترهای هندسی آنومالی سمت راست (عمیق‌تر) بهتر عمل کرده است البته میانگین خطا برای این مدل، به ازای تمام پارامترها برابر با ۴/۳۲ درصد می‌باشد که اشاره به دقت خوب این شبکه دارد.



شکل ۵-۱۶- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه PHASE (خط چین)

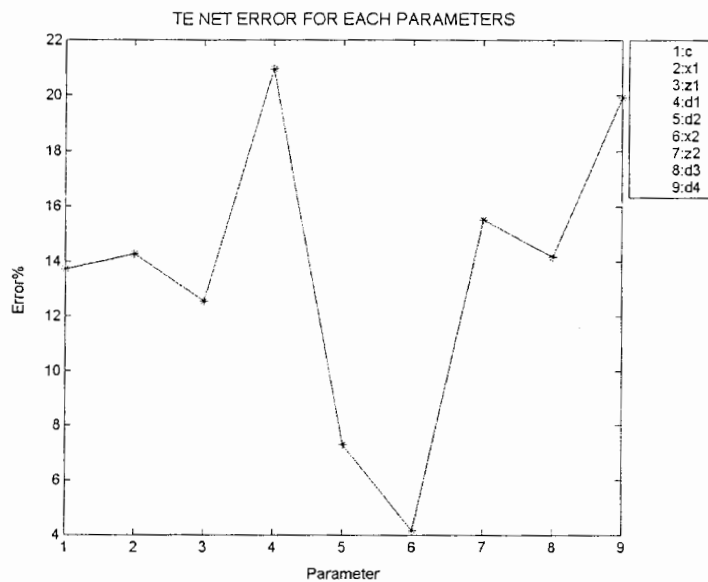
جدول ۵-۴- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه PHASE برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	55	51.664
N_1	2700	2640.3
Z_1	5700	5019.1
d_1	400	447.49
d_2	1200	1203.3
N_2	10150	9780.5
Z_2	13500	13454
d_3	1350	1338.9
d_4	4000	4068.1

۵-۸-۳- شبکه TE

این شبکه شامل دو پارامتر ورودی مقاومت ویژه ظاهری و فاز مد TE می‌باشد که به ترتیب به صورت ρ_{app} و ϕ_{app} که هر پارامتر برداری شامل ۹۹ مؤلفه می‌باشد، ساختار طراحی شده برای این شبکه ۹-۱۵-۱۴-۱۹۸ می‌باشد. مقدار خطای RMS به دست آمده در این شبکه برابر با ۰/۱۴۵ می‌باشد. میانگین خطای نسبی به دست آمده روی سری آزمایشی به ازای هر پارامتر در شکل ۵-۱۷ نشان داده شد و همچنین مقادیر عددی آن‌ها نیز به همراه شکل ارائه می‌شود.

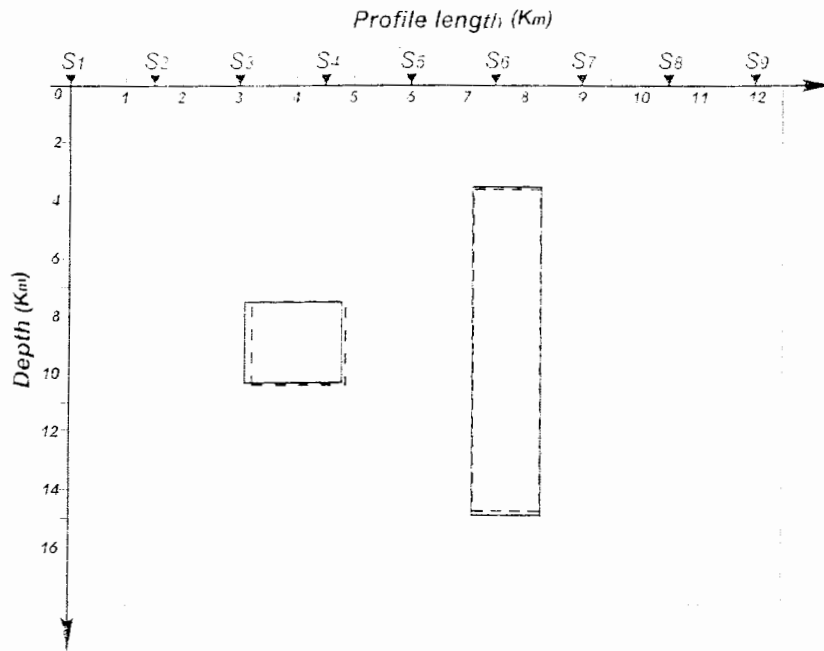
پارامتر	درصد خطا
c	13.726
x ₁	14.256
z ₁	12.552
d ₁	20.97
d ₂	7.2937
x ₂	4.1662
z ₂	15.506
d ₃	14.17
d ₄	19.918



شکل ۵-۱۷- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه TE به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ

با توجه به شکل و مقادیر مندرج در جدول همراه آن باز هم مشاهده می‌شود که مقدار خطای شبکه در مورد پارامترهای پنجم و ششم، کمینه می‌باشد و بیشینه آن در مورد پارامتر چهارم می‌باشد. در این شبکه خطای پارامتر ششم به حداقل مقدار نسبت به سایر شبکه‌ها می‌رسد. این امر ناشی از حساسیت شبکه‌ها نسبت به این پارامتر می‌باشد.

در ضمن برای مقایسه بهتر، مدل واقعی و مدل تخمینی توسط شبکه، به ازای یکی از نمونه‌های آزمایشی در شکل ۵-۱۸ نشان داده می‌شود، همچنین مقادیر عددی پارامترهای واقعی مدل به همراه پارامترهای تخمین زده شده توسط شبکه در جدول ۵-۵ ارائه شده است. که با توجه به شکل و نتایج مندرج در جدول ۵-۵ ملاحظه می‌شود این شبکه در تخمین پارامترهای مدل مخصوصاً در مورد آنومالی سمت راست بسیار دقیق عمل کرده است. این امر به دلیل بزرگ‌تر بودن آنومالی سمت راست می‌باشد که باعث ارائه مشخصه‌های بارزتر می‌گردد. میانگین خطای به دست آمده برای این مدل ۱/۸۸ درصد می‌باشد که بیانگر دقت بالای این شبکه می‌باشد.



شکل ۲-۱۸- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TE (خط چین)

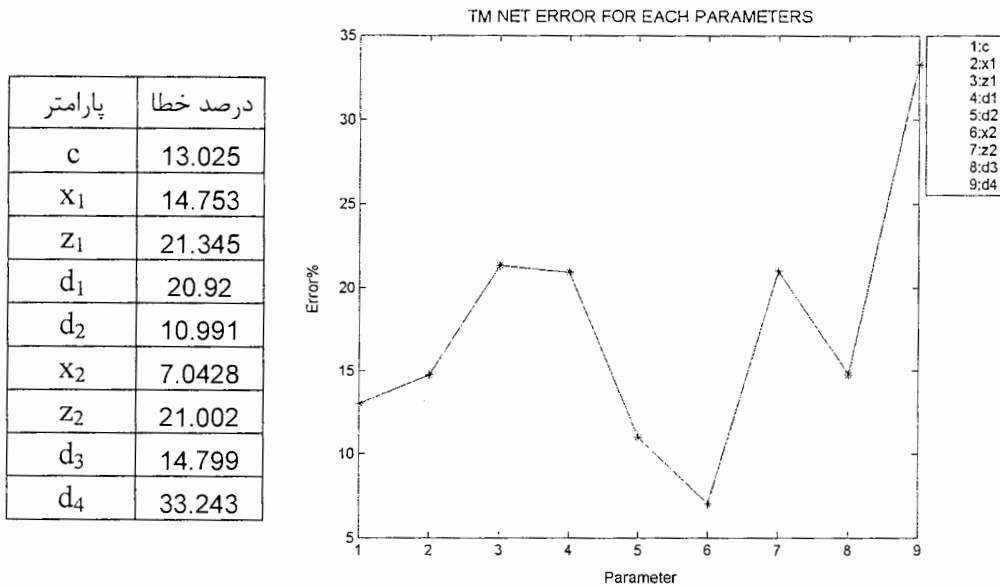
جدول ۲-۲- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TE برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل تخمین زده شده توسط شبکه	
	مدل واقعی	مدل تخمین زده شده
c	70	69.55
X_1	3950	4047.7
Z_1	8900	8937.5
d_1	850	819.44
d_2	1400	1451.2
X_2	7700	7692.4
Z_2	9200	9177.4
d_3	600	577.84
d_4	5700	5579.1

۸-۴- شبکه TM

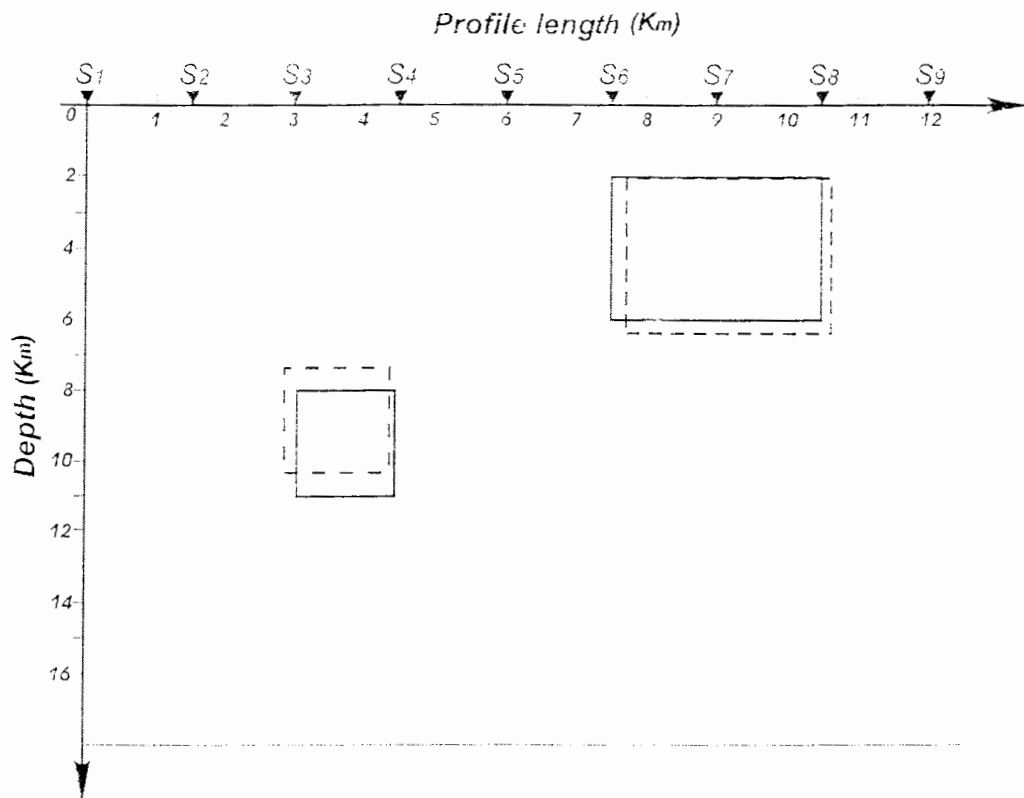
این شبکه شامل دو پارامتر ورودی ρ_{TM}^0 و φ_{TM}^0 به عنوان پارامترهای اندازه گیری شده مد TM می‌باشد و به همان ترتیب هر پارامتر برداری متشکل از ۹۹ مؤلفه است. ساختار شبکه به صورت ۹-۱۴-۱۹۸ و مقدار خطای RMS به دست آمده در این شبکه برابر با ۰/۱۷۸ می‌باشد. میانگین

خطای نسبی به دست آمده روی سری آزمایشی به ازای هر پارامتر در شکل ۵-۱۹ نشان داده می‌شود، و همچنین مقادیر عددی آن‌ها نیز در ستون همراه شکل ارائه می‌شود.



شکل ۵-۱۹- نمودار تغییرات میانگین خطای نسبی روی سری آزمایشی، به ازای هر پارامتر در شبکه TM به همراه مقادیر عددی آن‌ها ارائه شده در ستون سمت چپ

همان طور که ملاحظه می‌شود، کمینه خطا مربوط به پارامتر ششم می‌باشد و بیشینه خطا مربوط به پارامتر نهم می‌باشد. مقادیر خطا در این شبکه، در مورد بیشتر پارامترها نسبت به سایر شبکه‌های دیگر افزایش یافته است. برای ارزیابی صحت نتایج، مدل واقعی و مدل تخمین زده شده توسط این شبکه برای یکی نمونه‌های آزمایشی در شکل ۵-۲۰ نشان داده می‌شود، و همچنین مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TM در جدول ۵-۶ ارائه می‌شود، با توجه به شکل و نتایج مندرج در جدول ۵-۶، مشاهده می‌شود پارامترهای مربوط به آنومالی سمت راست (کم عمق) به طور دقیق‌تری نسبت به آنومالی سمت چپ (عمیق) تخمین زده شده است. دلیل این امر ناشی از حساسیت مد TM نسبت به آنومالی‌های کم عمق‌تر می‌باشد. میانگین خطای نسبی برای مدل مذکور ۴/۵۶ درصد می‌باشد.



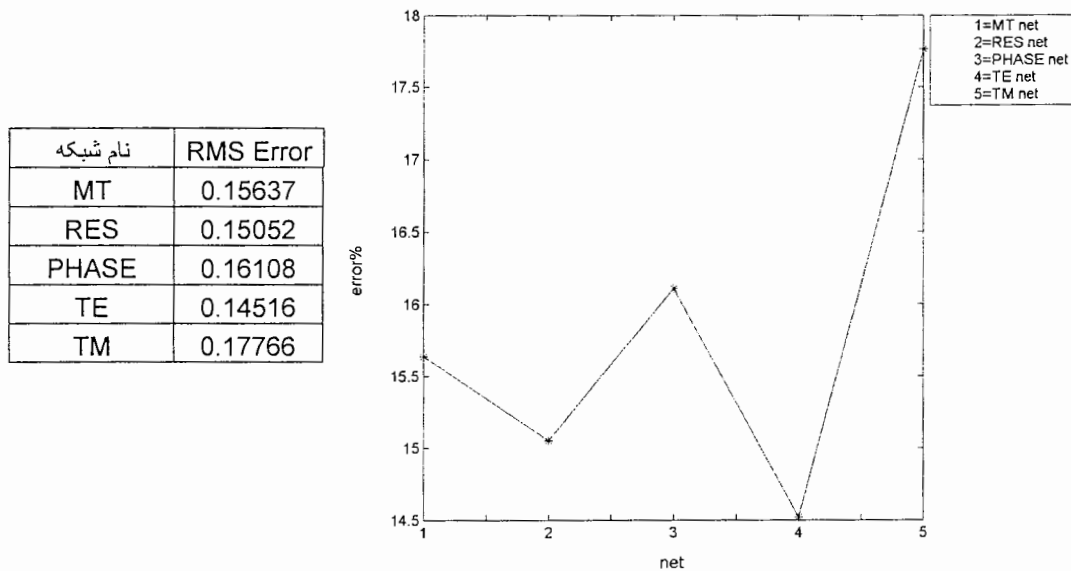
شکل ۵-۲۰- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TM (خط چین)

جدول ۵-۶- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TM برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل تخمین زده شده	
	مدل واقعی	توسط شبکه
c	40	35.36
N_1	3750	3626.7
Z_1	9500	8841.3
d_1	750	753.17
d_2	1500	1485.4
N_2	9000	9179
Z_2	4000	4188.7
d_3	1500	1463.3
d_4	2000	2173.9

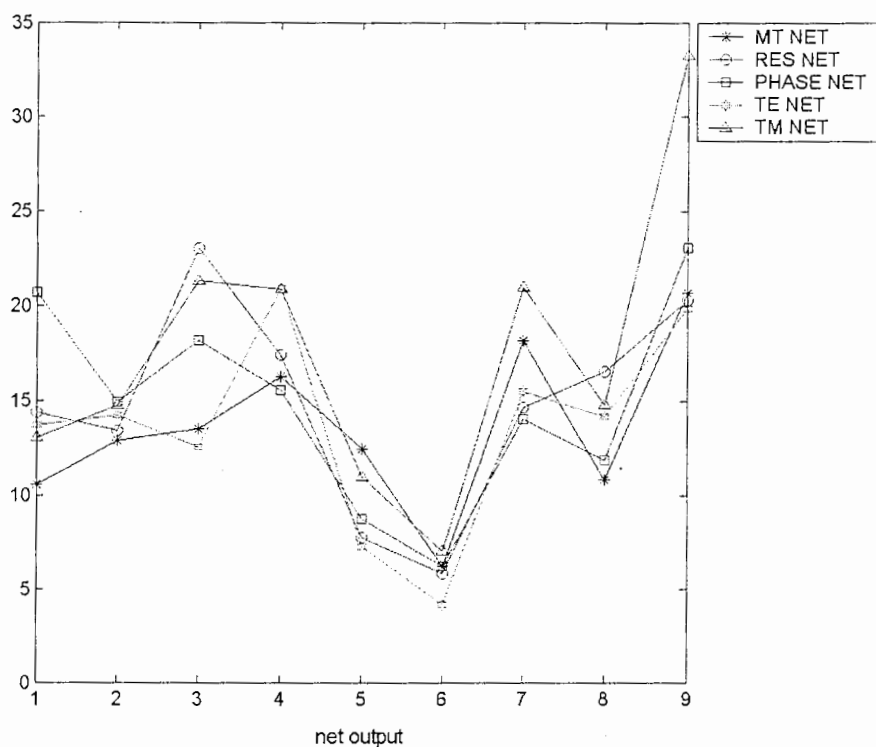
۵-۸-۵ - مقایسه و بررسی نتایج

به منظور مقایسه نتایج ۵ شبکه آموزش داده شده، خطای RMS به ازای هر شبکه در شکل ۵-۲۱ ارائه می‌شود، همچنین مقادیر عددی آن‌ها نیز در ستون همراه شکل ارائه می‌شود.



شکل ۵-۲۱- خطای RMS به ازای هر شبکه برای سری آزمایشی به همراه مقادیر عددی آن‌ها

ملاحظه می‌شود مقادیر خطا بسیار به هم نزدیک می‌باشد با این وجود، شبکه TE به کمترین خطای RMS حتی نسبت به شبکه MT دست می‌یابد. لازم به ذکر است که شبکه MT یک شبکه نسبتاً بزرگ است، از این رو به علت داشتن خطوط ارتباطی بیشتر و سعی در بالانس کردن تعداد وزن‌ها و بایاس‌های بیشتر، نیاز به تعداد مدل آموزشی بیشتری نیز دارد. از این رو با این تعداد مدل بدیهی است، خطا نیز افزایش می‌یابد. جهت بررسی و مقایسه بهتر، مقدار میانگین خطا به ازای هر پارامتر برای شبکه‌های مختلف در شکل ۵-۲۲ و مقادیر عددی آن‌ها در جدول ۵-۷ ارائه می‌شود.



شکل ۵-۲۲- مقادیر خطای میانگین سری آزمایشی به ازای هر پارامتر برای شبکه‌های مختلف

جدول ۵-۷- درصد خطای میانگین سری آزمایشی به ازای شبکه‌های مختلف

پارامتر	شبکه MT	شبکه RES	شبکه PHASE	شبکه TE	شبکه TM
c	10.546	14.362	20.723	13.726	13.025
x_1	12.917	13.406	14.939	14.256	14.753
z_1	13.507	23.035	18.18	12.552	21.345
d_1	16.307	17.419	15.572	20.97	20.92
d_2	12.45	7.7504	8.7351	7.2937	10.991
x_2	6.2278	5.8515	6.2194	4.1662	7.0428
z_2	18.2	14.68	14.022	15.506	21.002
d_3	10.859	16.562	11.881	14.17	14.799
d_4	20.712	20.341	23.119	19.918	33.243

ملاحظه می‌شود که در بیشتر پارامترها، شبکه TE نتیجه بهتری می‌دهد، علت را می‌بایست در

فضای مدل و خصوصیات مد TE جستجو کرد.

از آن جایی که بیشتر نمونه‌های آموزشی طراحی شده متعلق به مدل‌هایی می‌باشد که

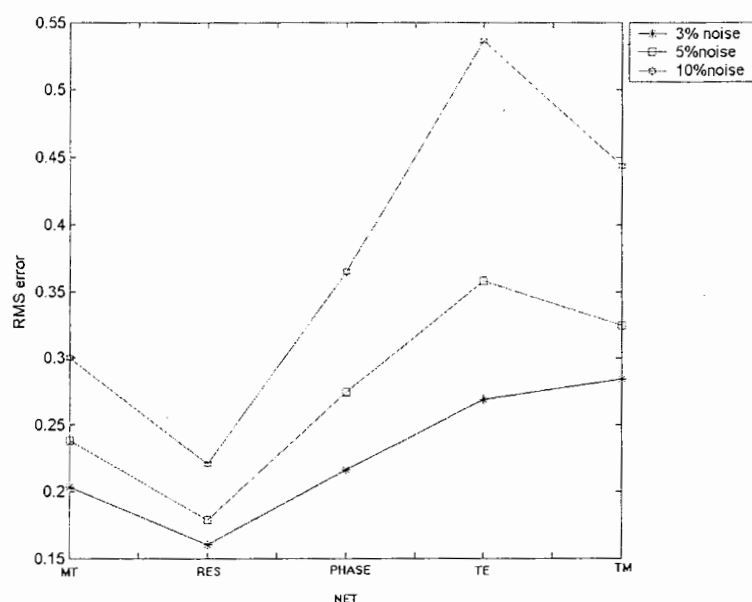
آنومالی‌ها در قسمت‌های عمیق قرار گرفته اند و مد TE معمولاً به آنومالی‌های عمیق حساس است، از

این رو با توجه به چگال بودن فضای مدل برای مد TE طبقاً این شبکه بعد از آموزش، پاسخ بهتری

می‌دهد. اما شبکه MII با وجود این که اطلاعات مربوط به مد TE را نیز در بر دارد نتیجه ضعیف‌تری نسبت به شبکه TE می‌دهد. علت این موضوع، بزرگ بودن شبکه MII می‌باشد از این رو این شبکه نیاز به سری آموزشی بزرگتری دارد. همان طور که ملاحظه می‌شود شبکه TM ضعیف‌ترین نتیجه را دارد. چون این مد بیشتر به آنومالی‌های کم عمق حساس می‌باشد. بنابراین به دلیل کم چگال بودن فضای مدل در این عمق، چنین نتایجی دور از انتظار نیست. در شکل‌های ارائه شده برای مدل‌های واقعی و تخمین زده شده توسط شبکه، به سهولت می‌توان میزان دقت و حساسیت دو شبکه TE و TM را در تخمین پارامترهای مدل مشاهده کرد. شبکه‌های RES و PHASE تقریباً بایک دقت، پارامترها را تخمین می‌زنند.

۵-۹- بررسی رفتار شبکه‌ها در حضور نویز

جهت بررسی رفتار شبکه‌ها در حضور نویز، مقادیر ۳، ۵ و ۱۰ درصد نویز با توزیع نرمال و میانگین صفر به طور تصادفی به داده‌های ورودی اضافه شد. علت اضافه کردن نویز به داده‌ها، شبیه‌سازی شرایط صحرایی و بررسی رفتار شبکه‌ها در برابر داده‌های صحرایی می‌باشد. در این حالت خطای RMS شبکه‌ها و میانگین خطای نسبی سری آزمایشی به ازای هر پارامتر بررسی شد. مقادیر خطای RMS به دست آمده برای هر شبکه و به ازای تزریق مقادیر ۳، ۵ و ۱۰ درصد نویز در شکل ۵-۲۳ نشان داده می‌شود. همچنین مقادیر عددی خطای RMS در حضور مقادیر مختلف نویز در جدول ۵-۸ ارائه می‌شود. با توجه به این مقادیر و نمودار رسم شده ملاحظه می‌شود با افزایش نویز، خطا نیز برای تمام شبکه‌ها افزایش می‌یابد اما میزان افزایش در هر شبکه متفاوت است، شبکه RES نسبت به شبکه‌های دیگر در حضور نویز مقاوم‌تر ولی شبکه TE ناپایدارتر نشان می‌دهد.



شکل ۵-۲۳- مقادیر خطای RMS به دست آمده برای شبکه‌های مختلف در حضور نویز

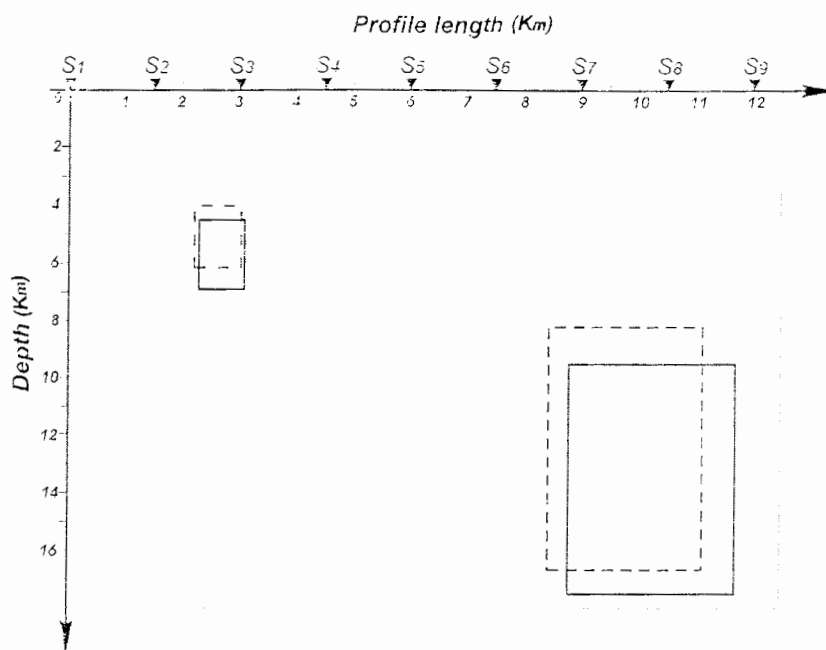
جدول ۵-۸- مقادیر RMS برای شبکه‌های مختلف در حضور نویز

نام شبکه	3%noise	5%noise	10%noise
MT	0.20328	0.23818	0.29994
RES	0.16012	0.17883	0.22067
PHASE	0.21629	0.2743	0.36502
TE	0.26906	0.35776	0.53609
TM	0.28412	0.32477	0.44336

در برداشت‌های صحرائی، حضور ۵ درصد نویز در داده‌ها معمول می‌باشد از این رو نمونه‌های انتخاب شده برای ارزیابی عملکرد شبکه‌ها در حضور نویز از نمونه‌هایی انتخاب شد که اولاً جزء سری آزمایشی بودند، به بیان دیگر این نمونه‌ها در فرآیند آموزش شرکت نداشتند و ثانیاً ۵ درصد نویز به آن‌ها تزریق شده بود.

به منظور درک بهتر از چگونگی تخمین پارامترها و ارزیابی شبکه‌ها در حضور نویز، ۵ مدل به عنوان نمونه، از سری آزمایشی انتخاب شد و پاسخ هر یک از شبکه‌ها در حضور ۵ درصد نویز، به همراه مدل واقعی تصویر شد. که مورد بررسی قرار می‌گیرد.

به همین منظور ابتدا پاسخ شبکه MT در حضور ۵ درصد نویز به ازای یکی از نمونه‌های آزمایشی در شکل ۵-۲۴ نشان داده می‌شود. همچنین مقادیر عددی پارامترهای واقعی و پارامترهای تخمین زده شده همین مدل در جدول ۵-۹ ارائه می‌شود.



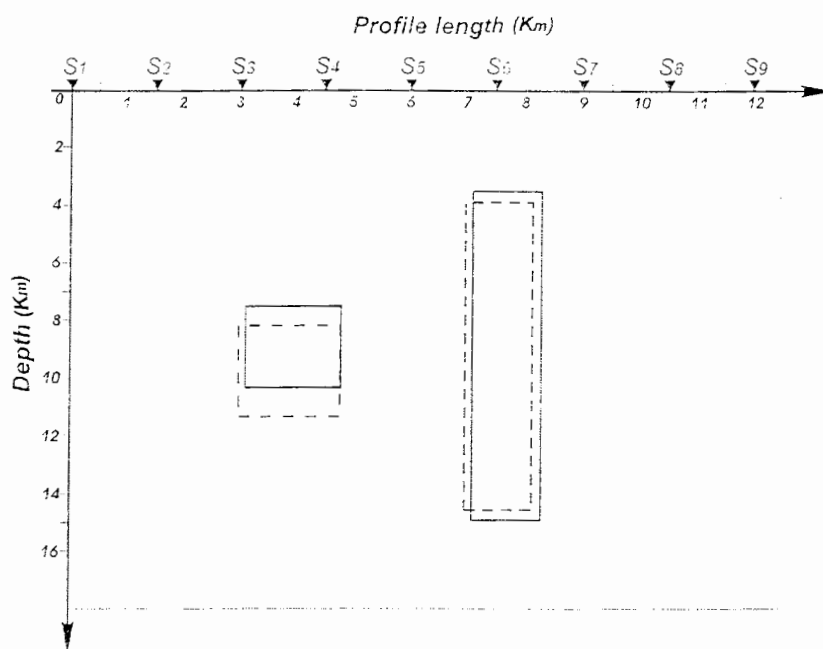
شکل ۵-۲۴- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه MT (خط چین) در حضور ۵ درصد نویز

جدول ۵-۹- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه MT برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	50	55.052
x_1	2700	2628.4
z_1	5700	5081.6
d_1	400	410.36
d_2	1200	1067.9
x_2	10250	9794.4
z_2	13500	12434
d_3	1450	1341.8
d_4	4000	4229

با توجه به شکل و مقادیر مندرج در جدول مربوطه مشاهده می‌شود، با وجود نویز شبکه MT توانسته پارامترهای مدل را با دقت قابل قبولی تخمین بزند این امر با محاسبه میانگین خطای ۶/۹۶

درصدی برای این مدل مشهود می‌باشد. در مورد شبکه RES نیز با ارائه یکی از نمونه‌های آزمایشی با تزریق ۵ درصد نویز نتایج زیر حاصل شد. این نتایج به صورت مدل واقعی و تخمینی توسط شبکه در شکل ۵-۲۵ تصویر شد، همچنین مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES در جدول ۵-۱۰ ارائه می‌شود. با توجه به شکل و نتایج به دست آمده ملاحظه می‌شود با وجود نویز این شبکه از دقت قابل قبولی برخوردار است به طوری که میانگین خطای نسبی برای این مدل برابر با ۵.۲۰۳ درصد می‌باشد

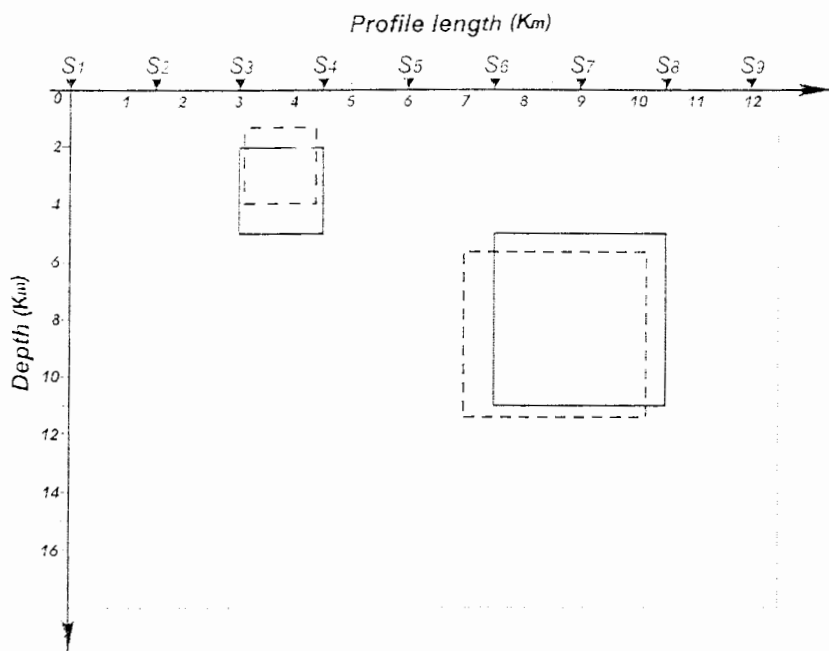


شکل ۵-۲۵- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه RES (خط چین) در حضور ۵ درصد نویز

جدول ۵-۱۰- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه RES برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	70	73.921
x_1	3950	3881.5
z_1	8900	9734.8
d_1	850	904.43
d_2	1400	1574.7
x_2	7700	7568.2
z_2	9200	9223.4
d_3	600	583.59
d_4	5700	5327.8

شبکه PHASE نیز با توجه به شکل ۵-۲۶ و نتایج مندرج در جدول ۵-۱۱ در حضور ۵ درصد نویز از دقت خوبی برخوردار می‌باشد به طوری که میانگین خطای نسبی برای این مدل ۸,۲۶۶ درصد می‌باشد.



شکل ۵-۲۶- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه PHASE (خط چین) در حضور ۵ درصد نویز

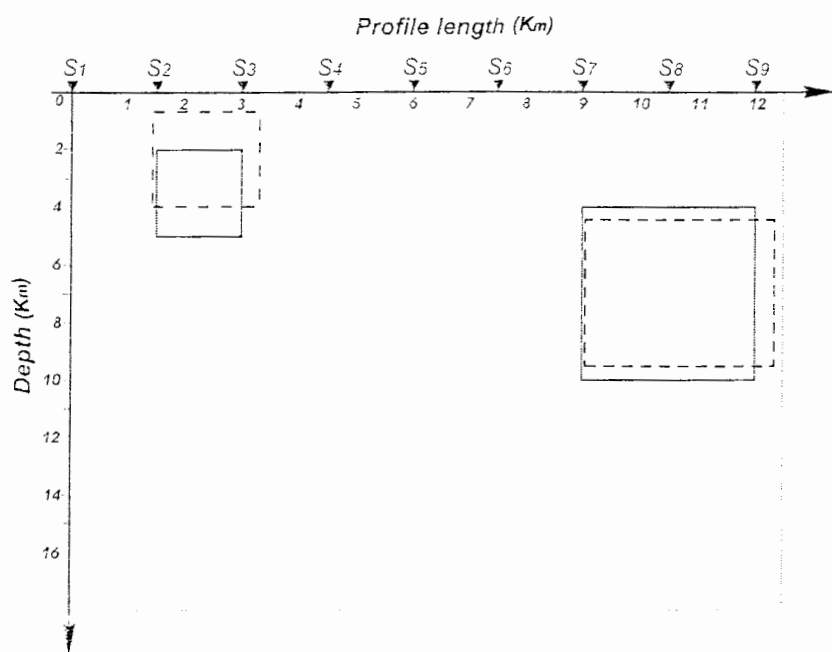
جدول ۵-۱۱- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه PHASE برای یکی از نمونه‌های

آزمایشی

پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	100	99.163
x_1	3750	3734.7
z_1	3500	2621.2
d_1	750	641.87
d_2	1500	1324.6
x_2	9000	8564
z_2	8000	8527.7
d_3	1500	1594.9
d_4	3000	2875.1

رفتار شبکه TE در برابر ورودی دارای ۵ درصد نویز نسبت به شبکه‌های دیگر دارای خطای بیشتری می‌باشد. این امر به وضوح در نمونه آزمایشی ارائه شده در شکل ۵-۲۷ نشان داده می‌شود.

همچنین مقادیر عددی پارامترهای واقعی مدل و تخمین زده شده توسط شبکه در جدول ۵-۱۲ ارائه می‌شود. میانگین خطای نسبی برای این مدل ۱۱/۸۸۹ درصد می‌باشد که نسبت به سایر شبکه‌ها بیشتر می‌باشد.



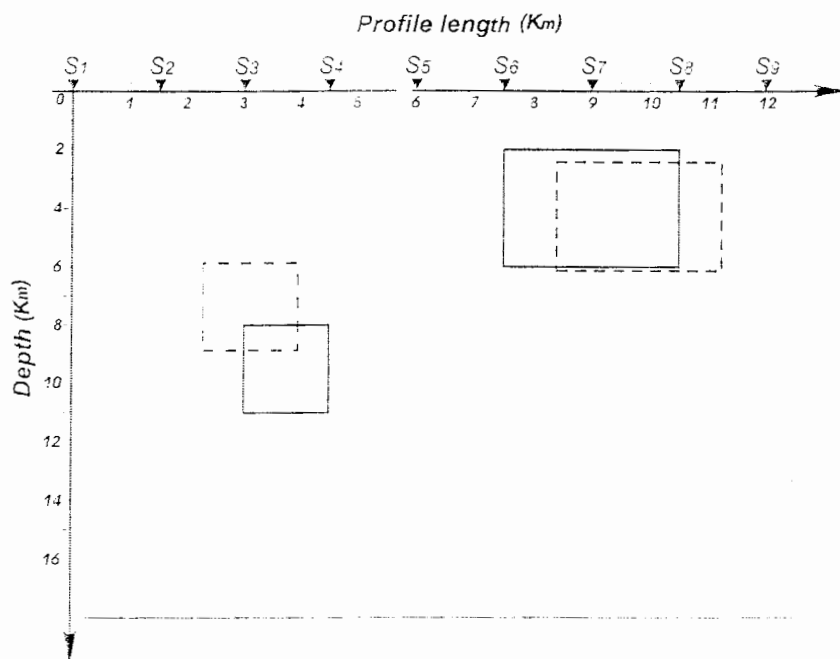
شکل ۵-۲۷- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TE (خط چین) در حضور ۵ درصد نویز

جدول ۵-۱۲- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TE برای یکی از نمونه‌های آزمایشی

پارامترهای مدل	مدل تخمین زده شده	
	مدل واقعی	توسط شبکه
c	80	84.488
x ₁	2250	2373.6
z ₁	3500	2321.4
d ₁	750	940.65
d ₂	1500	1649
x ₂	10500	10696
z ₂	7000	6977
d ₃	1500	1638.2
d ₄	3000	2536.2

همان طور که قبلاً اشاره شد و در شکل بالا نیز مشاهده می‌شود شبکه TE پارامترهای مربوط به آنومالی عمیق‌تر را دقیق‌تر تخمین می‌زند چون مد TE نسبت به آنومالی‌های عمیق حساستر

می‌باشد. پاسخ شبکه TM نیز در حضور ۵ درصد نویز به ازای یکی از نمونه‌های آزمایشی در شکل ۵-۲۸ تصویر شده است. مقادیر عددی پارامترهای واقعی و پارامترهای تخمین زده شده مربوط به این مدل در جدول ۵-۱۳ ارائه می‌شود.



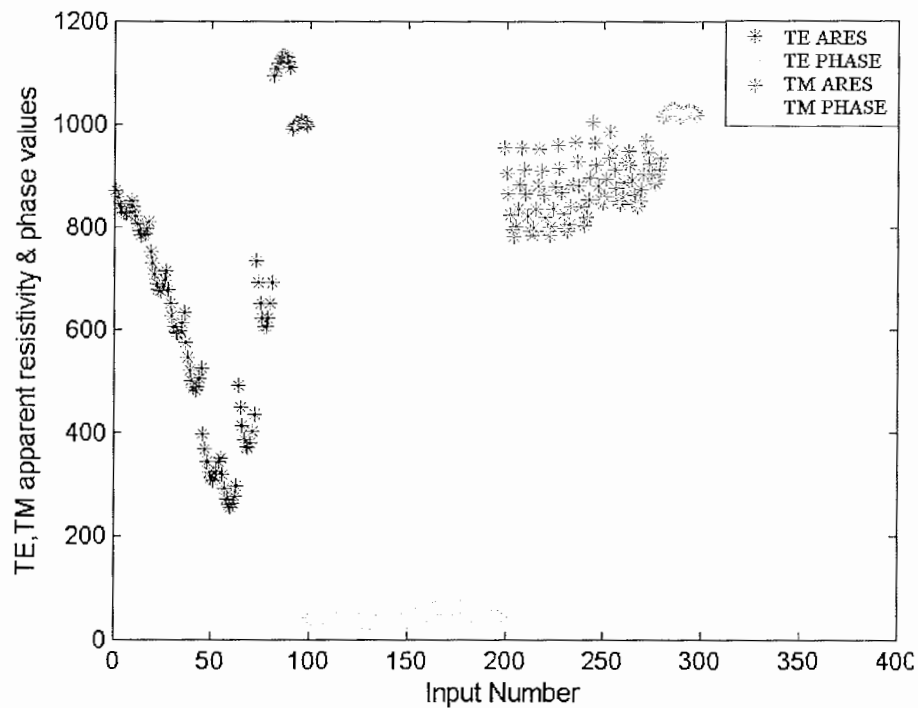
شکل ۵-۲۸- مدل واقعی (خط پر) و مدل تخمینی توسط شبکه TM (خط چین) در حضور ۵ درصد نویز

جدول ۵-۱۳- مقادیر عددی پارامترهای واقعی و تخمین زده شده توسط شبکه TM برای یکی از نمونه‌های آزمایشی

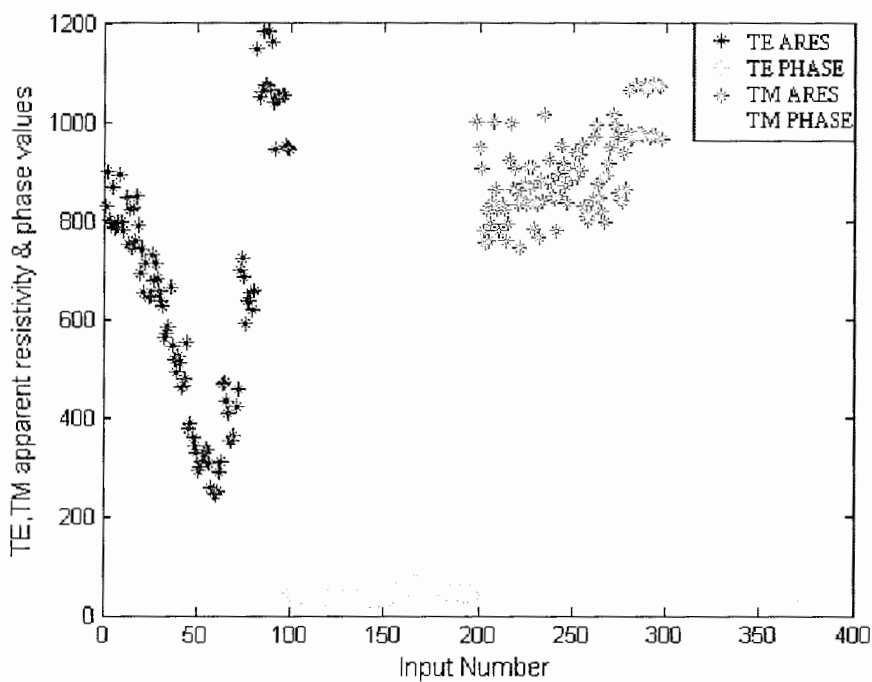
پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
c	40	43.846
x_1	3750	3111.1
z_1	9500	7368.8
d_1	750	840.4
d_2	1500	1499.7
x_2	9000	9822.4
z_2	4000	4281.3
d_3	1500	1420.9
d_4	2000	1857.7

همان طور که در شکل ۵-۲۸ ملاحظه می‌شود این شبکه پارامترهای آنومالی کم عمق‌تر را با دقت بیشتری تخمین می‌زند این امر به دلیل حساسیت مد TM به آنومالی‌های کم عمق می‌باشد. میانگین خطای نسبی برای این مدل ۹/۹۶ درصد می‌باشد. که این خطا بیشتر ناشی از عدم توانایی این شبکه در تخمین پارامترهای آنومالی عمیق‌تر می‌باشد.

همان طور که بیان شد در حضور نویز، به علت مغشوش شدن برخی از مشخصه‌هایی که شبکه روی آن‌ها آموزش دیده است، خطا افزایش می‌یابد. به عنوان مثال تغییرات مقادیر مقاومت ویژه و فاز را در هر دو مد برای یکی از مدل‌های آزمایشی در غیاب نویز و در حضور ۵ درصد نویز بررسی می‌کنیم. به همین منظور ۹۲ امین مدل یا ۹۲ امین ستون از ماتریس ورودی را انتخاب کردیم که از سری نمونه‌های آزمایشی می‌باشد. در شکل ۵-۲۹ نمودار این مدل در غیاب نویز نشان داده می‌شود. این نمودار چگونگی تغییرات مقاومت ویژه و فاز به ترتیب برای دو مد TE و TM را نشان می‌دهد. شبکه در خلال آموزش این تغییرات را به عنوان مشخصه‌های ورودی به ازای یک خروجی معین یاد می‌گیرد، البته چگونگی یادگیری این مشخصه‌ها به تعداد مدل‌های آموزشی، ساختار شبکه و روش تعمیم آن نیز بستگی دارد. لازم به ذکر است این مشخصه‌ها با توجه به محل آنومالی‌ها، اندازه آن‌ها و تباین مقاومت ویژه آنومالی‌ها و زمینه ایجاد می‌شوند. با اضافه کردن مقداری نویز به ورودی‌ها وضوح برخی از این مشخصه‌ها از بین می‌رود (شکل ۵-۳۰) اما روند کلی نمودار تغییر نکرده است. از این رو خطای شبکه به منظور تخمین پارامترها کمی افزایش می‌یابد ولی به دلیل عدم تغییر در روند کلی تغییرات مقاومت ویژه و فاز در دو مد هنوز شبکه از دقت قابل قبولی در تخمین پارامترهای مدل برخوردار می‌باشد.



شکل ۵-۲۹- تغییرات مقادیر مقاومت ویژه و فاز در دو مد TE و TM، مربوط به یکی از نمونه‌های آزمایشی در غیاب نویز



شکل ۵-۳۰- تغییرات مقادیر مقاومت ویژه و فاز در دو مد TE و TM، مربوط به یکی از نمونه‌های آزمایشی در حضور ۵ درصد نویز

به منظور بررسی کیفیت آموزش در تخمین پارامترهای مدل و اعتبارسنجی شبکه، شبه مقاطع خطای نسبی مقاومت ویژه و فاز، بین مدل واقعی و مدل تخمین زده شده توسط شبکه MT در هر دو مد TE و TM، در غیاب نویز (شکل ۵-۳۱) و در حضور ۵ درصد نویز (شکل ۵-۳۲)، برای یکی از مدل‌های آزمایشی، به عنوان نمونه رسم شده است که مورد بحث قرار می‌گیرد. به همین منظور با ارائه مقادیر مقاومت ویژه و فاز مدهای TE و TM به صورت بدون نویز و حاوی ۵ درصد نویز، مربوط به یکی از نمونه‌های آزمایشی به شبکه MT، پاسخ شبکه به صورت پارامترهای مدل دریافت می‌شود. سپس با پارامترهای مدل مذکور مدلی ساخته می‌شود و سپس پاسخ آن به صورت مقادیر مقاومت ویژه و فاز هر مد با انجام مدل‌سازی مستقیم توسط نرم‌افزار EMIXMT2D ارائه می‌شود. خطای نسبی مقادیر مقاومت ویژه و فاز مدهای TE و TM در هر فرکانس به صورت نظیر به نظیر، محاسبه می‌شود که به صورت شبه مقطع در شکل‌های ۵-۳۱ (بدون نویز) و ۵-۳۲ (در حضور ۵ درصد نویز) ارائه می‌شود. همان‌طور که در شکل ۵-۳۱ مشاهده می‌شود مقدار خطا در شبه مقاطع مقاومت ویژه مد TE بین ۴ تا ۱۰ درصد و در مد TM بین ۸ تا ۱۰ درصد می‌باشد، همچنین مقدار خطا در شبه مقاطع فاز برای مدهای TE و TM در بیشتر نقاط حول ۱ تا ۳ درصد در غیاب نویز می‌باشد، با توجه به مقادیر تخمین زده شده (جدول ۵-۲) این امر نشان می‌دهد که شبکه به خوبی آموزش داده شده است و تعمیم خوبی یافته است. برای بررسی بهتر مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT در غیاب نویز در جدول ۵-۱۴ ارائه می‌شود.

جدول ۵-۱۴- مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT

آنومالی	پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
	(Ωm) مقاومت ویژه زمینه	500	539.5
سمت چپ	(m) محل مرز سمت چپ	2300	2283.84
	(m) محل مرز سمت راست	3100	3099.76
	(m) محل مرز بالایی	4500	4593.8
	(m) محل مرز پایینی	6900	6741.8
سمت راست	(m) محل مرز سمت چپ	8800	8310.9
	(m) محل مرز سمت راست	11700	11008.1
	(m) محل مرز بالایی	9500	9384
	(m) محل مرز پایینی	17500	17508

با توجه به مقادیر مندرج در جدول ۵-۱۴ ملاحظه می‌شود مرزهای آنومالی‌ها که توسط شبکه MT تخمین زده شده اند از دقت بسیار خوبی برخوردار است.

و همچنین همان طور که در شکل ۵-۳۲ مشاهده می‌شود مقدار خطا در شبه مقاطع مقاومت ویژه مد TE بین ۲ تا ۱۰ درصد و در مد TM نیز به همین صورت می‌باشد و تنها در تعداد محدودی از نقاط خطا افزایش می‌یابد. و همچنین مقدار خطا در شبه مقطع فاز برای مد TE حدود ۲ تا ۸ درصد و مد TM نیز در حدود ۴ تا ۶ درصد در حضور نویز می‌باشد، با توجه به مقادیر تخمین زده شده (جدول ۵-۹) این امر نشان می‌دهد که عملکرد شبکه در حضور نویز نیز دارای دقت قابل قبولی می‌باشد.

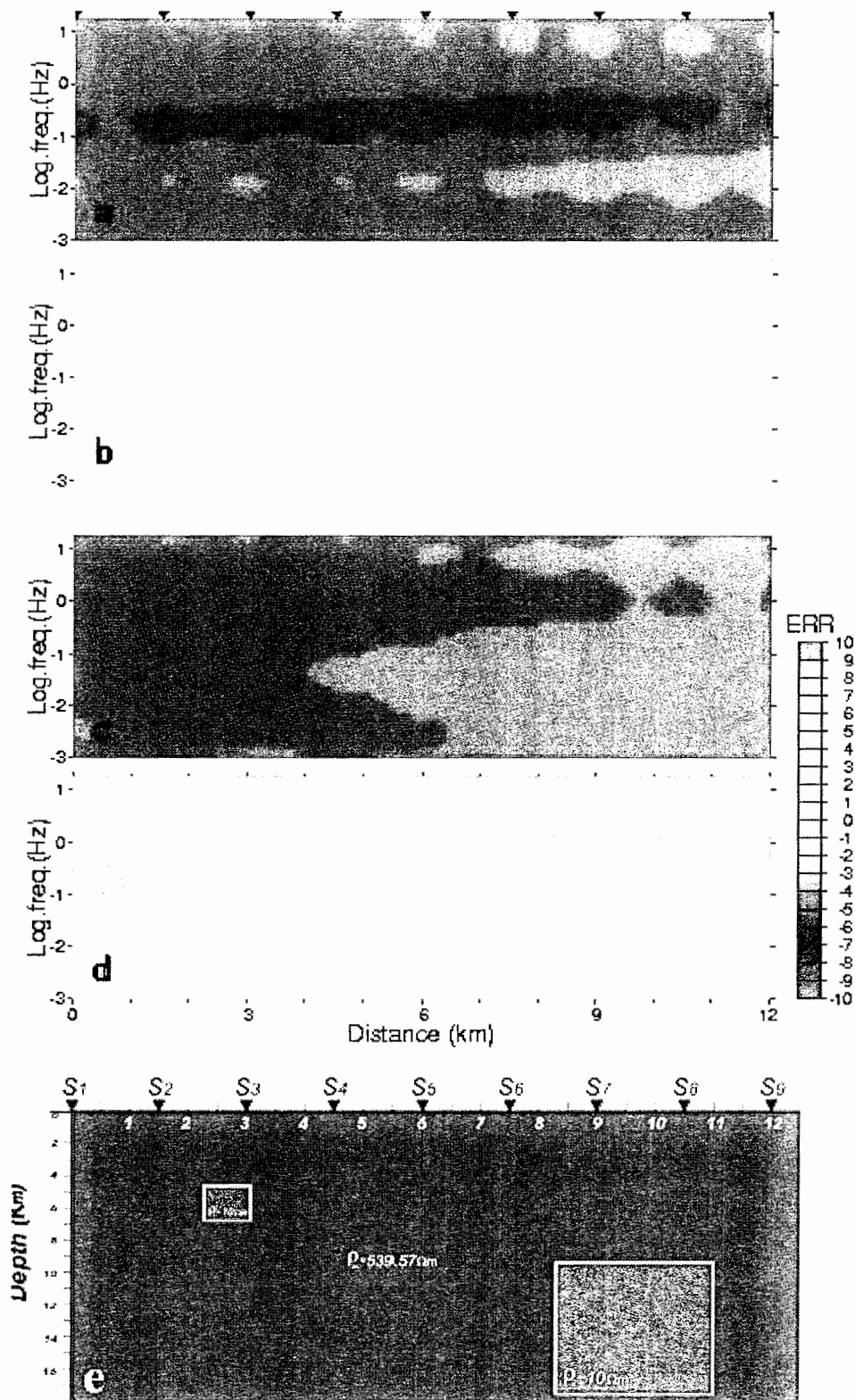
برای بررسی بهتر مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT در حضور ۵ درصد نویز در جدول ۵-۱۵ ارائه می‌شود.

جدول ۵-۱۵- مقادیر واقعی و تخمین زده شده مقاومت ویژه زمینه و مرزهای آنومالی توسط شبکه MT با ۵ درصد

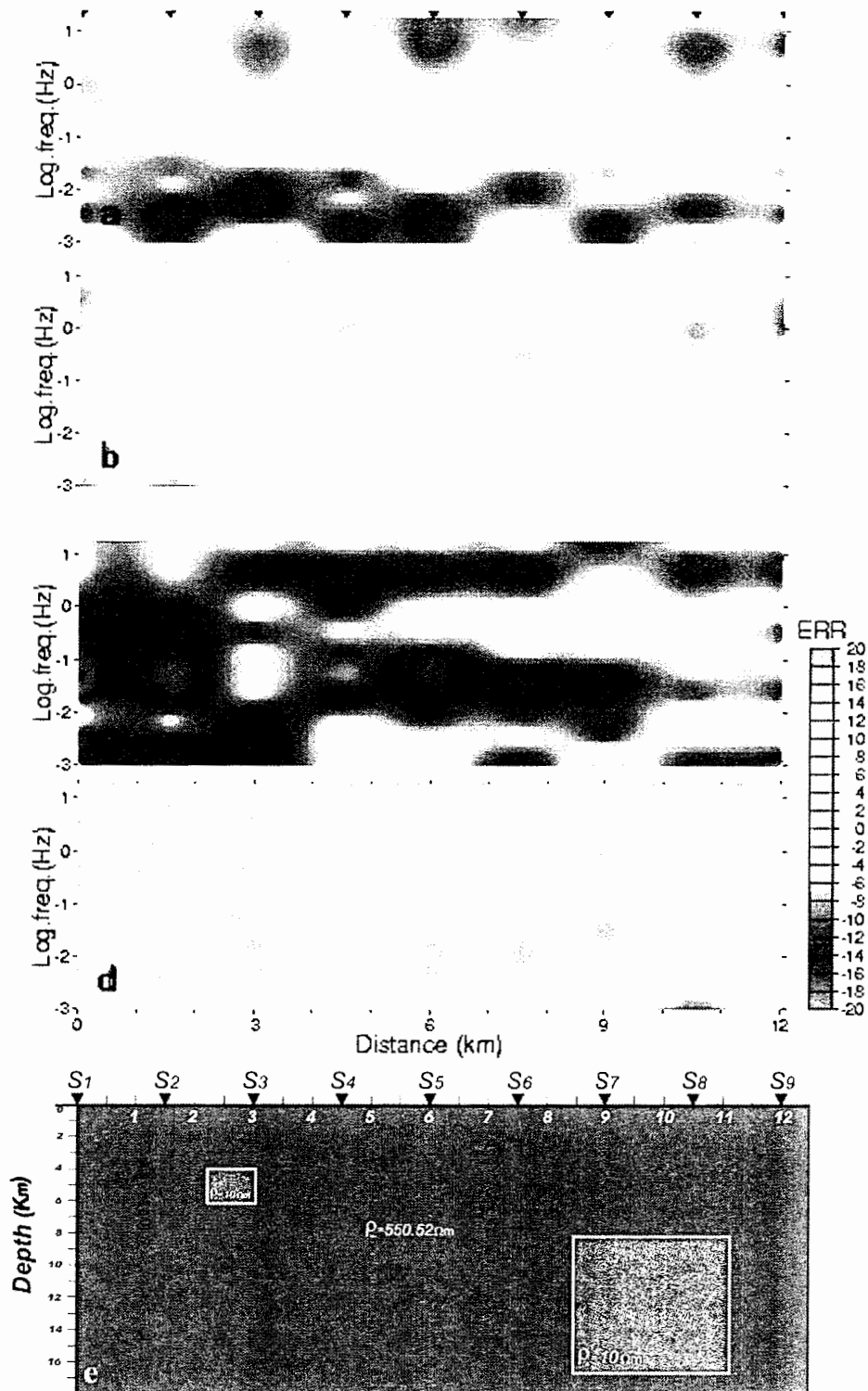
نویز

آنومالی	پارامترهای مدل	مدل واقعی	مدل تخمین زده شده توسط شبکه
	مقاومت ویژه زمینه (Ωm)	500	550.52
سمت چپ	محل مرز سمت چپ (m)	2300	2218.04
	محل مرز سمت راست (m)	3100	3038.76
	محل مرز بالایی (m)	4500	4013.7
	محل مرز پایینی (m)	6900	6149.5
سمت راست	محل مرز سمت چپ (m)	8800	8452.6
	محل مرز سمت راست (m)	11700	11136.2
	محل مرز بالایی (m)	9500	8205
	محل مرز پایینی (m)	17500	16663

با توجه به مقادیر ارائه شده در جدول ۵-۱۵ مشاهده می‌شود که شبکه آموزش داده شده در حضور ۵ درصد نویز نیز با دقت قابل قبولی پارامترها را تخمین می‌زند. داده‌های حاوی حدود ۵ درصد نویز می‌تواند تقریبی از داده‌های صحرائی باشد.



شکل ۵-۳۱- (a)، شبه مقطع خطای نسبی مقاومت ویژه مد TE، (b) شبه مقطع خطای نسبی فاز مد TE، (c) شبه مقطع خطای نسبی مقاومت ویژه مد TM، (d) مدل تخمین زده شده توسط شبکه MT در غیاب نویز، (e) شبه مقطع خطای نسبی فاز مد TM



شکل ۵-۳۲- شبه مقطع خطای نسبی مقاومت ویژه مد TE (a)، شبه مقطع خطای نسبی فاز مد TE (b)، شبه مقطع خطای نسبی مقاومت ویژه مد TM (c)، شبه مقطع خطای نسبی فاز مد TM (d)، مدل تخمین زده شده توسط شبکه MT (e) در حضور ۵ درصد نویز

۵-۱۰- آموزش شبکه برای تخمین یک پارامتر

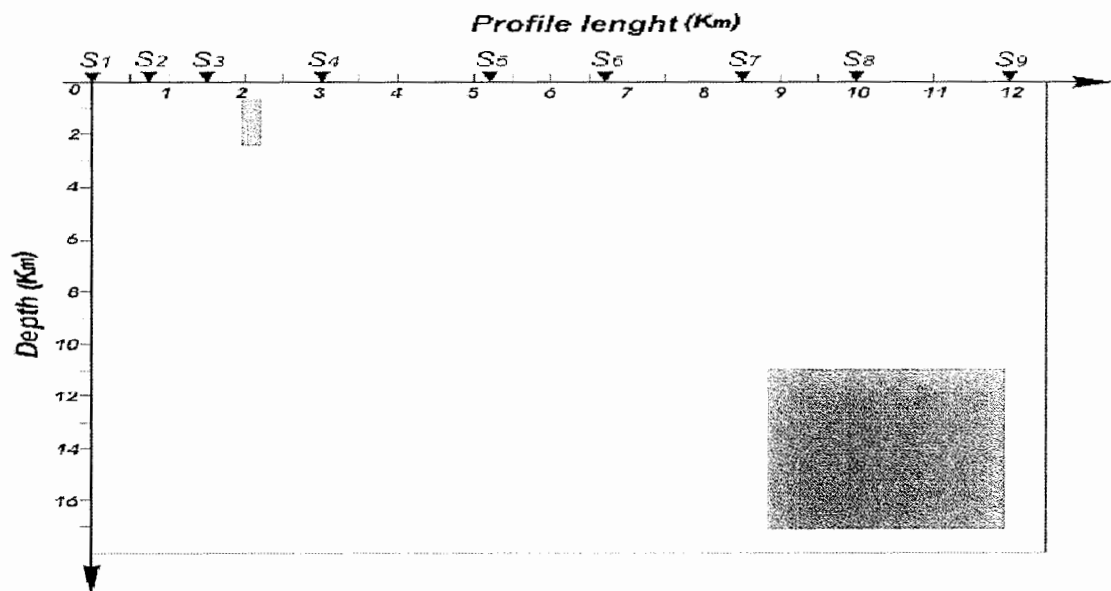
چنانچه شبکه به منظور تخمین تنها یک پارامتر آموزش ببیند، بدیابی است مقدار خطای تخمین پارامتر کاهش می‌یابد، برای نمونه شبکه‌ای برای تخمین پارامتر اول آموزش داده شد. مقدار خطای میانگین حاصل از شبکه آموزش داده شده برای تخمین پارامتر c برابر با $4/17$ درصد، می‌باشد در صورتی که میانگین خطای به دست آمده برای تخمین پارامتر c با شبکه MT برابر با $10/54$ درصد، می‌باشد. لازم به ذکر است که به دلیل محدودیت زمانی، تعداد نرون‌های لایه پنهان برای این شبکه به بیش از ۶ نرون افزایش نیافت و احتمال می‌رود که با افزایش تعداد نرون‌ها خطا نیز بیشتر کاهش یابد.

۵-۱۱- آزمایش شبکه با داده‌های صحرائی

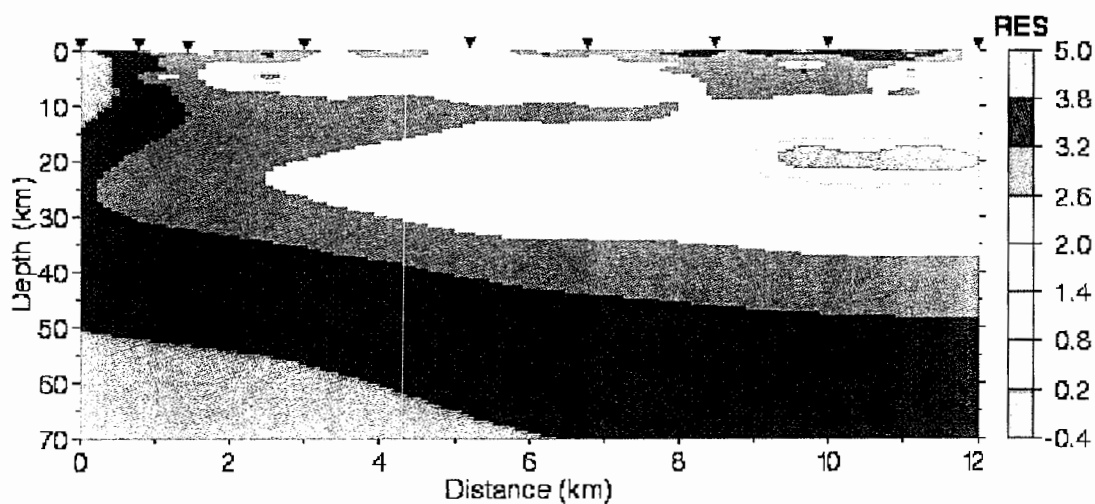
به منظور بررسی عملکرد شبکه در برابر داده‌های صحرائی، وارونه‌سازی این داده‌ها توسط شبکه MT صورت گرفت که نتایج آن در جدول ۵-۱۶ ارائه می‌شود. مدلی بر اساس پارامترهای تخمین زده شده توسط شبکه رسم شد که در شکل ۵-۳۳ نشان داده می‌شود، همچنین مدل‌سازی هموار ترکیبی دو مدل TE و TM این داده‌ها با روش RRI صورت گرفت. شکل ۵-۳۴ مقطع حاصل از این مدل‌سازی که توسط نرم افزار گرافیکی GMT رسم شده است را نشان می‌دهد.

جدول ۵-۱۶- تخمین پارامترهای داده‌های صحرائی توسط شبکه MT

پارامترهای مدل	پاسخ شبکه
c	95.484
x_1	2089.1
z_1	1489.5
d_1	121.75
d_2	871.01
x_2	10400
z_2	14023
d_3	1527.7
d_4	3064



شکل ۵-۳۳- مدل تخمین زده شده برای داده‌های صحرائی با شبکه MT



شکل ۵-۳۴- مقطع مقاومت ویژه حاصل از وارونه سازی هموار ترکیبی دو مد TE و TM [مرادزاده، ۱۹۹۸]

همان‌طور که مشاهده می‌شود تخمین محل آنومالی‌ها توسط شبکه با محل ارائه شده در مقطع مدل‌سازی هموار همخوانی دارد. با توجه به این که مدل‌سازی هموار قادر به تخمین دقیق مرزهای آنومالی‌ها نیست و حدود آنومالی‌ها به صورت هاله مشخص می‌شود، مشاهده می‌شود ابعاد حدودی آنومالی‌ها شامل بعد عرضی آنومالی سمت چپ، ابعاد طولی و عرضی آنومالی سمت راست با مقادیر تخمین زده شده این پارامترها، توسط شبکه MT همخوانی دارد. ولی مقادیر عمقی آنومالی‌ها و بعد طولی آنومالی سمت چپ با مقادیر تخمین زده شده توسط شبکه همخوانی ندارد. همچنین روش

مدل‌سازی هموار اجازه می‌دهد تا مقاومت ویژه زمینه تغییر تدریجی داشته باشد که با توجه به مقطع ارائه شده در شکل ۵-۳۴ مشاهده می‌شود گستره این تغییرات در اغلب نقاط بین ۸۰۰ تا ۱۰۰۰ اهم متر می‌باشد. در صورتی که در نمونه‌های آموزشی مورد استفاده در شبکه مقدار مقاومت ویژه، ثابت در نظر گرفته شده است. با این وجود مقدار پارامتر تخمین زده شده مقاومت ویژه توسط شبکه ۹۵۴/۸۵ اهم متر می‌باشد که در گستره تعیین شده در مقطع مدل‌سازی هموار قرار دارد، این امر بیانگر میزان دقت شبکه می‌باشد. لازم به ذکر است با اطلاعات موجود، بدون استفاده از اطلاعات اضافی دیگر و یا پیش از حفاری نمی‌توان در مورد این که کدامیک از این دو روش دقیق‌تر می‌باشند، نظر داد. چون مدل ارائه شده در شکل ۵-۳۴ بر اساس روش مدل‌سازی هموار به دست آمده است و همان‌طور که قبلاً ذکر شد این روش مدل‌سازی قادر به تخمین دقیق محل مرزها نمی‌باشد. از سوی دیگر مدل واقعی زمین، قاعدتاً بسیار پیچیده‌تر از کلاس مدل در نظر گرفته شده، می‌باشد بنابراین چنانچه شکل هندسی مفروض برای آنومالی‌ها به جای مستطیل، بیضی یا چند ضلعی انتخاب می‌شد شاید به دلیل نزدیک شدن به مدل واقعی زمین، نتایج بهتری حاصل می‌شد.

پس از بررسی عملکرد شبکه‌های آموزش داده شده، جمع بندی و نتایج به دست آمده به همراه چند پیشنهاد در فصل بعد ارائه می‌شود.

فصل ششم:

جمع بندی، نتیجه گیری و پیشنهادات

جمع بندی، نتیجه گیری و پیشنهادات

۱-۶- جمع بندی

با توجه به مشکلات ذکر شده در مورد مدل سازی وارون داده های مگنتوتلوریک، حجیم بودن داده ها، وقت گیر بودن وارونه سازی با استفاده از روش های معمول و کم دقت بودن این روش ها در تخمین پارامترها، ضرورت استفاده و آزمایش روش های جدید در وارونه سازی داده های مگنتوتلوریک کاملاً محسوس می باشد. با توجه به توانمندی های شبکه های عصبی به ویژه در تخمین توابع، به کار گیری این روش به عنوان یکی از روش های حل مشکلات فوق در وارونه سازی داده های مگنتوتلوریک به عنوان هدف این پایان نامه تعیین شد. به منظور دسترسی به این هدف ابتدا با توجه به شبه مقاطع مقاومت ویژه و فاز داده های صحرائی، کلاس مدل انتخاب شد. استفاده هم زمان داده های مقاومت ویژه و فاز در دو مد و چندین فرکانس و ارائه این داده ها به شبکه، منجر به بزرگ شدن اندازه شبکه می شد که طراحی و آموزش آن را با توجه به امکانات کامپیوتری در دسترس بسیار مشکل می نمود. از این رو مجبور به حذف داده های چند فرکانس شدیم که این امر باعث کاهش تفکیک پذیری شد. از سوی دیگر تولید داده ها در ۱۱ فرکانس و ۹ ایستگاه باعث شد که مثلاً در شبکه MT تعداد ورودی های شبکه، ۳۹۶ داده برای هر مدل شود، که پس از آموزش و طراحی این شبکه به دلیل این که ماتریس های وزنی و بردارهای بایاس ذخیره شده با این حالت تطبیق یافتند لازم است هر الگوی ارائه شده به شبکه جهت تخمین پارامترهای مدل دارای ۳۹۶ داده مقاومت ویژه و فاز به ترتیب در مدهای TE و TM باشد. البته حل این معضل به کمک روش های کاهش اندازه مانند نمونه گیری امکان پذیر می باشد.

به منظور آموزش شبکه و تولید نمونه‌های آموزشی از نرم‌افزاری استفاده شد که این نرم‌افزار برای تولید داده‌ها از یکی از روش‌های مدل‌سازی مستقیم استفاده می‌کرد. هر چند روش شبکه عصبی و روش‌های معمول وارونه‌سازی هر دو بر پایه استفاده از روش‌های مدل‌سازی مستقیم می‌باشند ولی روش شبکه عصبی به شیوه متفاوتی از این مدل‌سازی استفاده می‌کند.

در نهایت و با وجود مشکلات طراحی و آموزش شبکه‌های بزرگ، این شبکه‌ها طراحی شد. به طوری که بررسی نتایج به دست آمده روی نمونه‌ها آزمایشی نشان می‌دهد که شبکه‌های آموزش داده شده قادر به بازسازی مدل زیر سطحی با دقت قابل قبولی در غیاب و یا حضور نویز می‌باشند. همچنین بررسی و مقایسه نتایج به دست آمده با استفاده از شبکه‌های عصبی برای یک مجموعه داده‌های صحرائی و نتایج حاصل از وارونه‌سازی این داده‌ها با استفاده از یکی از روش‌های معمول وارونه‌سازی (RRI) نشان می‌دهد که با وجود انتخاب کلاس ساده‌ای برای داده‌های صحرائی، این دو روش در تخمین بسیاری از پارامترهای مدل زمینی با یکدیگر همخوانی دارند. این امر نشان می‌دهد که شبکه‌ها به خوبی طراحی، آموزش و تعمیم یافتند. ماحصل این تحقیق دستیابی به چند نتیجه و ارائه چند پیشنهاد جهت بهبود کار می‌باشد که در ذیل به آن پرداخته می‌شود.

۲-۶- نتیجه‌گیری

۱. مزیت اصلی استفاده از شبکه‌های عصبی در وارونه‌سازی داده‌ها این است که شبکه یک بار در کلاس خاصی از مسائل آموزش می‌بیند و چنانچه آموزش و تعمیم شبکه به خوبی صورت گیرد. شبکه قادر به تخمین خروجی برای یک ورودی معین در مدت زمان بسیار کوتاه می‌باشد.
۲. با توجه به امکانات کامپیوتری محدود، به منظور کاهش تعداد ورودی‌ها به شبکه ما مجبور به حذف داده‌های برخی از فرکانس‌ها شدیم که این امر باعث کاهش تفکیک پذیری و متعاقباً کاهش دقت شبکه شد.
۳. مهمترین بخش در آموزش یک شبکه، تولید نمونه‌های آموزشی مناسب می‌باشد. نکته مهم در تولید نمونه‌های آموزشی، حفظ تباین خوب بین مقاومت ویژه زمینه و آنومالی‌ها با توجه به ابعاد آنومالی‌ها، عمق آن‌ها و فرکانس‌های اندازه‌گیری می‌باشد.

۴. شرط اصلی برای دستیابی شبکه به خطای کمتر، در اختیار داشتن تعداد بیشتری از داده‌های مناسب می‌باشد. هر چه تعداد نمونه‌های آموزشی با خروجی متفاوت، اما ورودی نزدیک به هم در بین داده‌ها بیشتر باشد شبکه دچار خطای بیشتری می‌شود.
۵. شبکه TE در تخمین برخی از پارامترها، نتیجه بهتری حتی نسبت به شبکه MT می‌دهد، به دلیل این‌که بیشتر نمونه‌های آموزشی تولید شده در قسمت‌های عمیق زمین قرار گرفته‌اند و مد TE به آنومالی‌های عمیق حساس است.
۶. شبکه MT با وجود داشتن داده‌های مد TE در ماتریس ورودی، در تخمین برخی از پارامترها نسبت به شبکه TE ضعیف عمل می‌کند، این امر به این علت است که شبکه MT یک شبکه نسبتاً بزرگ است، بنابراین جهت ایجاد توازن بین وزن‌ها و نمونه‌های آموزشی به سری آموزشی بزرگتری نیز نیاز دارد.
۷. شبکه TM بیشترین خطا را در تخمین برخی از پارامترهای مدل نسبت به سایر شبکه‌ها دارد، این امر به دلیل عدم پوشش مناسب قسمت‌های کم عمق زمین است، چون شبکه TM نسبت به آنومالی‌های کم عمق حساس است.
۸. شبکه‌های RES و PHASE در تخمین پارامترهای مجهول مدل تقریباً با دقت یکسانی عمل می‌کنند.
۹. در بررسی رفتار شبکه‌ها در تخمین پارامترهای مدل در حضور نویز، مشخص شد که خطا در تمام شبکه‌ها با افزایش نویز، افزایش می‌یابد اما میزان افزایش خطا در هر شبکه متفاوت است. به طوری که شبکه RES در حضور نویز مقاوم و شبکه TE ناپایدار نشان می‌دهد.
۱۰. چنانچه به جای آموزش یک شبکه برای تخمین ۹ پارامتر مجهول مدل، از شبکه‌های مجزا برای تخمین هر پارامتر استفاده شود دقت مدل‌سازی به طور محسوس افزایش می‌یابد.
۱۱. در مقایسه نتایج شبکه به ازای داده‌های صحرائی و نتایج حاصل از یکی از روش‌های معمول مدل‌سازی وارون، مشاهده می‌شود تخمین پارامترهای مقاومت ویژه زمینه، بعد عرضی آنومالی سمت چپ و ابعاد طولی و عرضی آنومالی سمت راست در هر دو روش همخوانی خوبی دارد و تنها عمق آنومالی‌ها و بعد طولی آنومالی سمت چپ در دو روش اختلاف دارند.

۱۲. در مورد این که در تخمین پارامترهای مدل استفاده از کدام روش (شبکه عصبی و روش معمول مدل سازی وارون) بر روی داده‌های صحرایی دقیق‌تر می باشد، پیش از حفاری و یا استفاده از اطلاعات اضافی دیگر نمی‌توان نظری داد.

۳-۶- پیشنهادات

۱. به منظور افزایش دقت شبکه، حذف نمونه‌های آموزشی نامناسب و پالایش داده‌ها پیشنهاد می‌شود.
۲. آموزش و طراحی یک شبکه بزرگ و تخمین هم زمان ۹ پارامتر کار بسیار مشکلی می‌باشد و همان طور که مشاهده شد آموزش شبکه‌های جداگانه برای تخمین هر پارامتر میزان دقت شبکه را در تخمین پارامترها افزایش می‌دهد ولی از آن جایی که آموزش و طراحی ۹ شبکه کار وقت گیری می‌باشد. آموزش ۳ شبکه برای تخمین ۳ پارامتر به طور هم زمان پیشنهاد می‌شود.
۳. برای آموزش شبکه از یک مدل بسیار ساده استفاده شده است در صورتی که ساختار ژئوالکتریکی می‌تواند بسیار پیچیده باشد. شکل هندسی انتخاب شده برای آنومالی‌ها بسیار ساده و به صورت مستطیل در نظر گرفته شده است. چنانچه به جای استفاده از این مدل از اشکال هندسی پیچیده‌تر مانند، هشت ضلعی و یا بیضی استفاده شود مدل فرضی به مدل واقعی نزدیکتر می‌شود.
۴. به منظور ساده شدن مسأله، رسانایی آنومالی‌ها برابر با هم و ثابت فرض شد، این مورد ممکن است در طبیعت کمتر دیده شود. ترکیب نمونه‌های آموزشی متنوع‌تر در همین کلاس می‌تواند منجر به توانمندتر کردن شبکه شود.
۵. شبکه آموزش داده شده، تنها روی داده‌های متعلق به کلاس آشنا برای شبکه پاسخ مناسب می‌دهد، بنابراین کلاس داده‌های مشاهده‌ای می‌باید تعیین شود. از این رو طراحی و استفاده از یک شبکه دسته‌بندی کننده برای تعیین کلاس داده‌ها پیشنهاد می‌شود.
۶. در بررسی روش مگنتوتلوریک به داده‌های tipper به عنوان یکی از داده‌هایی که برای تفسیر استفاده می‌شود، نیز اشاره شد. اثر حضور این داده‌ها در ماتریس ورودی به علت محدودیت

زمانی بررسی نشد از این رو اضافه کردن این داده‌ها در ماتریس ورودی شبکه و بررسی اثر آن در تخمین پارامترها پیشنهاد می‌شود.

۷. در این شبکه‌ها تابع فعالیت انتخاب شده برای نرون‌های هر لایه یکسان فرض شده است. چون شبکه برای تخمین چند پارامتر به کار می‌رود ممکن است انواع دیگری از توابع فعالیت برای تخمین هر پارامتر مناسب باشد، از این رو انتخاب و آزمایش تابع فعالیت‌های متفاوت برای نرون‌های مختلف پیشنهاد می‌شود.

۸. تولید داده‌های در تعداد بیشتری از فرکانس‌ها و ارائه آن‌ها به شبکه، می‌تواند باعث افزایش تفکیک پذیری شود، البته این امر باعث افزایش تعداد ورودی‌ها به شبکه و متعاقباً بزرگ شدن اندازه شبکه می‌شود که در این وضعیت استفاده از روش‌های کاهش اندازه پیشنهاد می‌شود.

۹. کاهش اندازه، به عنوان یک روش بسیار مؤثر در تعیین دقت شبکه‌های عصبی می‌باشد. شاید تمام اطلاعات ارائه شده به عنوان ورودی، جهت آموزش آن لازم نباشد. فشرده‌سازی داده‌ها با استفاده از برخی از تبدیل‌ها و یا نمونه‌گیری داده‌ها می‌تواند، یک شبکه کوچکتر ایجاد کند، که نیاز به سری آموزشی کمتری نیز دارد. فشرده‌سازی اطلاعات باید به گونه‌ای انجام شود که غنی بودن داده‌ها از بین نرود، در غیر این صورت نه تنها دقت افزایش نمی‌یابد بلکه به دلیل حذف اطلاعات مهم کاهش نیز می‌یابد. از این رو استفاده از روش‌های کاهش اندازه مانند استفاده از تبدیل فوریه پیشنهاد می‌شود.

۱۰. برای تعمیم این شبکه روش منظم‌سازی مورد آزمایش قرار نگرفت. شاید این روش برای داده‌های ما مناسب‌تر باشد چون استفاده از روش توقف سریع، تعداد نمونه‌های آموزشی را کاهش می‌دهد. از این رو آزمایش روش منظم‌سازی با استفاده از کامپیوترهای سریع و یا با کاهش تعداد داده‌های ورودی، پیشنهاد می‌شود. حتی می‌توان از ترکیب این دو روش تعمیم دهی، نیز استفاده کرد.

فهرست منابع

فهرست منابع

انبرزی، م.، ۱۳۸۰. آشنایی با شبکه‌های عصبی مصنوعی. موسسه انتشارات علمی دانشگاه صنعتی شریف، تهران.

بخشی، ا.، ۱۳۸۲. شبیه‌سازی نمودار شاخص فتوالکترونیک با شبکه‌های عصبی مصنوعی. پایان نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود، دانشکده معدن و ژئوفیزیک.

ریتس، ج.ر.، میلفورد، ف.ج. و کریستی، ر.، ۱۳۷۴. مبانی نظریه الکترومغناطیس. صمیمی، ج.، جمشیدی پور، و عزیزاده قمصری، ن. (ترجمه). چاپ چهارم، مرکز نشر دانشگاهی، تهران.

گرامی زادگان، ا.، ۱۳۸۱. مدل‌سازی عددی در ژئوفیزیک. سمینار کارشناسی ارشد، دانشگاه صنعتی شاهرود، دانشکده معدن و ژئوفیزیک.

گرامی زادگان، ا.، ۱۳۸۲. معکوس سازی دوبعدی داده‌های ژئوالکترونیک با استفاده از شبکه‌های عصبی مصنوعی. پایان نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود، دانشکده معدن و ژئوفیزیک.

نوبار، ال.، مورینو، ا. و فنلون، ژ.پ.، ۱۳۷۴. پردازش و تحلیل داده‌های آماری روش‌ها و برنامه‌ها. توتونیان، ف. و صادقی، ح. (ترجمه). چاپ اول، انتشارات دانشگاه فردوسی مشهد.

مرادزاده، ع.، عرب امیری، ع.، ۱۳۸۳. مدل‌سازی معکوس پارامتری داده‌های پلاریزاسیون القایی و مقاومت ویژه اندیس معدنی پاینده. هشتمین همایش انجمن زمین شناسی ایران. شاهرود.

منیاج، د.ب.. ۱۳۷۹. مبانی شبکه‌های عصبی مصنوعی. جلد اول و دوم؛ انتشارات دانشگاه صنعتی امیرکبیر.

نقی زاده، د. ۱۳۸۲. کاربرد روش مگنتو تلوریک در اکتشاف ذخایر ژئو ترمال منطقه خوی. پایان نامه کارشناسی ارشد، دانشگاه تهران، مؤسسه ژئوفیزیک.

Asik, U., Gunel, T. and Erer, I., Neural network approach to the extraction of information about conducting cylinders from amplitude-only data. <http://www.eleco.emo.org.tr/eleco2003/oralB.htm-13x>.

Bailey, S.A., Ocean color model inversion using an artificial neural network. <http://www.obs1.wff.nasa.gov/bailey/Neural1>.

Baum, E.B. and Hausler, D., 1989. What size network gives valid generalization? *Neural Comp*, **1**, 151-160.

Bescoby, D.J., Cawley, G.C. and Chroston, P.N., 2004. Interpretation of Geophysical Surveys of Archeological Sites using Artificial Neural Networks. <http://bioinfo.cpgei.cefetpr.br/anais/IJCNN2003/Papers/1204-333>.

Bishop, C., 1995. *Neural Networks for Pattern Recognition*. Oxford Press.

Calderon-Macias, C., Sen, M.K. and Stoffa, P.L., 1997. Hopfield neural networks, and mean field annealing for seismic deconvolution and multiple attenuation. *Geophysics*, **62**, 992-1002.

Calderon-Macias, C., Sen, M.K. and Stoffa, P.L., 1998. Automatic NMO correction and velocity estimation by a feedforward neural network. *Geophysics*, **63**, 1696-1707.

Calderon-Macias, C., Sen, M.K. and Stoffa, P.L., 2000. Artificial neural networks for parameter estimation in geophysics. *Geophysical Prospecting*, **48**, 21-47.

Cantwell, T., 1960. *Detection and analysis of low frequency magnetotelluric signals*. Mass.

Inst. Tech.

Cary, P. and Upham, W., 1992. An evaluation of neural networks. *The leading edge of exploration*, 45-47.

Constable, S.C., Parker, R.L. and Constable, C.G., 1987. Occams' inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**, 289-300.

Dai, H., Li, X.Y. and MacBeth, C., 1999. Inversion of Fracture Parameters by Using the Artificial Neural Network. *IC-AI'99 International Conference*, 517-523. Edinburgh.

de Groot-Heldin, C. and Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics*, **55**, 1613-1624.

de Groot-Heldin, C. and Constable, C., 2004. inversion of magnetotelluric data for 2D structure with sharp resistivity contrasts. *Geophysics*, **69**, 78-86.

de Lugao, P.P. and Wannamaker, P.E., 1996. Calculating the two-dimensional magnetotelluric Jacobian in finite elements using reciprocity. *Geophysical Journal International*, **127**, 806-810.

Dobrin, M.B. and Sovit, C.H., 1988. Introduction to Geophysical prospecting. McGraw-Hill, Inc., 868.

Dowla, F.U., Taylor, S.R. and Anderson, R.W., 1990. Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data. *Bull. Seis. Soc. Am*, **80**, 1346-1373.

El-Qady, G. and Ushijima, K., 2001. Inversion of DC resistivity data using neural networks. *Geophysical Prospecting*, **49**, 417-430.

Geman, S., Bienestock, E. and Doursat, R., 1992. Neural networks and bias/variance dilemma. *Neural Computation*, **4**, 1-58.

Gosse, C. and Siessons, B., 2000. Lithology prediction in the Taranaki Basin using neural network applications. *2000 New Zealand Petroleum Conference Proceedings*. New Zealand.

Griffiths, D.H. and King, R.F., 1988. Applied Geophysics for Geologists and Engineers, 230.

Hecht-Nielsen, R., 1987. Kolmogorov's mapping and neural network existence theorem. *1st Internat. Con. on Neural networks, IEEE Proc*, 11-14.

Hidalgo, H., Gomez-Trevino, E. and Swiniarski, R., 1994. Neural network approximation of an inverse functional. *IEEE World Congress on Neural Networks and Fuzzy Systems*. Orlando.

Hoerl, A.E. and Kennard, R.W., 1970a. Ridge regression: Based estimation for nonorthogonal problems. *Thechnometric*, **12**, 55-67.

Hoerl, A.E. and Kennard, R.W., 1970b. Ridge regression: Applications to nonorthogonal problems. *Technometrics*, **12**, 69-82.

Huang, S.C. and Huang, Y.F., 1991. Bounds on the number of hidden neurons in multilayer perceptrons. *IEEE Trans. Neur. Networks*, **2**, 47-55.

Huang, Z., Shimeld, J., Williamson, M. and Katsube, J., 1996. Permeability prediction with artificial neural network modeling in the Ventura gas field, offshore eastern Canada. *Geophysics*, **61**, 422-436.

Hush, D., Horne, B. and Salas, J.M., 1992. Error surface for multilayer perceptrons. *IEEE Trans. Systems*, **22**, 1152-1161.

Hush, D.R. and Horne, B.G., 1993. Progress in Supervised Neural Networks. *IEEE SIGNAL PROCESSING MAGAZINE*, 8-38.

Inman, J.R., Ryo, J. and Ward, S.H., 1973. Resistivity inversion. *Geophysics*, **38**, 1088-

1108.

Inman, J.R., 1975. Resistivity inversion with ridge regression. *Geophysics*, **40**, 798-87.

Jupp, D.L.B. and Vozoff, K., 1977. Two-dimensional magnetotelluric inversion. *Geophysical Journal of the Royal Astronomical Society*, **50**, 333-352.

Kaufman, A.A. and Keller, G.V., 1981. The magnetotelluric sounding method. Elsevier, 595.

Langer, R.E., 1933. An inversion problem in differential equation. *Am. Math. Soc*, **29**, 814-820.

Langer, H., Nunnari, G. and Occhipinti, L., 1996. Estimation of seismic waveform governing parameters with neural networks. *J. Geophys. Res*, **101**, 20109-20118.

Lippmann, R., 1987. An Introduction to Computing with Neural Nets. *IEEE ASSP Mag*, 417-434.

Loke, M.H. and Barker, R.D., 1996. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method. *Geophysical Prospecting*, **44**, 131-152.

Mackie, R.L., Madden, T.R. and Wannamaker, P.E., 1993. Three-dimensional magnetotelluric modeling using difference equation-Theory and comparisons to integral equations. *Geophysics*, **58**, 215-226.

Manry, M.T., Apollo, S.J. and Yu, Q., 1996. Minimum mean square estimation and neural networks. *Neurocomputing*, **13**, 59-74.

Marcuello-Pascual, A., Kaikkonen, P. and Pous, J., 1992. 2-D Inversion of MT data with a variable model geometry. *Geophysical Journal International*, **110**, 297-304.

Martf, M., 2001. Neural Network Prediction in a System for Optimizing Simulations. www.bus.colorado.edu/faculty/Laguna/articles/nnpred.

Mehanee, S. and Zhdanov, M., 2002. Two-dimensional magnetotelluric inversion of blocky geoelectrical structures. *Journal of Geophysical Research*, **107**, 1-11.

Melsa, P.J.W., 1989. Neural Networks: A Conceptual Overview. Tellabs Research Center.

Moradzadeh, A., 1998. Electrical imaging of the Adelaide geosyncline using magnetotelluric (MT). Ph.D Thesis, Flinders University of south Australia, 334.

Moradzadeh, A., 2003. using of tipper function to map subsurface conductivity structures in magnetotelluric (MT) surveys. *21st Geoscience Conference and Exposition*. Tehran.

Moradzadeh, A., 2003. Static shift appraisal and its correction in magnetotelluric (MT) surveys. *21st Geoscience Conference and Exposition*. Tehran.

Oristaglio, M.L. and Worthington, M.H., 1980. Inversion of surface and borehole electromagnetic data for two-dimensional electrical conductivity models. *Geophysical Prospecting*, **28**, 633-657.

Papilinski, A.P., 2002. Feedforward Multilayer Neural Networks. www.csse.monash.edu.au/~app/CSE5301/Lnts/L05.

Poulton, M.M., Sternberg, B.K. and Glass, C.E., 1992. Location of subsurface targets in geophysical data using neural networks. *Geophysics*, **57**, 1534-1544.

Poulton, M.M., Sternberg, B.K. and Glass, C.E., 1992. Neural network pattern recognition of subsurface EM images. *Journal of Applied Geophysics*, **29**, 21-36.

Poulton, M.M., 2001. Computational neural networks for geophysical data processing. PERGAMON, 335.

Raiche, A., 1991. A pattern recognition approach to geophysical inversion using neural nets. *Geophys. J. Int*, **105**, 629-648.

Rodi, W., 1989. Regularization and Backus-Gilbert estimation in nonlinear inverse

problems: Application to magnetotelluric and surface waves. Ph.D Thesis, Pennsylvania State University.

Rodi, W. and Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion. *Geophysics*, **66**, 174-187.

Romeo, G., 1994. Seismic signals detection and classification using artificial neural networks. *Annali di Geofisica*, **37**, 343-353.

Roth, G. and Tarantola, A., 1994. Neural networks and inversion of seismic data. *Journal of geophysical Research*, **99**, 6753-6768.

Smith, D.R., 1974. Variational methods in optimization. Prentice-Hall, Inc.

Smith, J.T. and Booker, J.R., 1988. Magnetotelluric inversion for minimum structure. *Geophysics*, **53**, 1565-1576.

Smith, J.T. and Booker, J.R., 1991. Rapid inversion of two- and three- dimensional magnetotelluric data. *J. Geophys. Res.*, **96**, 3905-3922.

Smith, T., Hoversten, M., Gasperikova, E. and Morrison, F., 1999. sharp boundary inversion of 2D magnetotelluric data. *Geophysical Prospecting*, **47**, 469-486.

Soulie, F., Gallinari, P., Le, C. and Thiria, S., 1987. Evaluation of neural network architectures on test learning tasks. *1st Internat. Conf. on Neural Networks*, 663-670.

Spichak, V. and Popova, I., 1998. Application of the neural network approach to the reconstruction of a three-dimensional geoelectric structure. *Physics of the Solid Earth*, **34**, 33-39.

Spichak, V. and Popova, I., 2000. Artificial neural network inversion of magnetotelluric data in terms of three- dimensional earth macroparameters. *Geophys. J. Int.*, **142**, 15-26.

Spichak, V., Fukuoka, K., Kobayashi, T., Mogi, T., Popova, I. and Shima, H., 2002. ANN

reconstruction of geoelectrical parameters of the Minou fault zone by scalar CSAMT data. *Journal of Applied Geophysics*, **49**, 75-90.

Stoyer, C., Zerelli, A., Slee, C. and Buttler, M.S., 1999. Manual of EMIXMT2D. Interpex.

Telford, W.M., Geldart, L.P. and Sheriff, R.E., 1990. Applied Geophysics. Cambridge University Press, 770.

The MathWork, I., 2002. Matlab, The language of technical computing. The MathWork, Inc.

Tikhonov, A.N., 1963a. Regularization of incorrectly posed problems. *Soviet Math. Dokl*, **4**, 1035-1038.

Tikhonov, A.N., 1963b. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl*, **4**, 1624-1627.

Tillery, S. and Combs, N., 1987. The problem of internal representations in neural nets: concepts, implications, and a proposed metric. *1st Internat. Conf. on Neural Networks, IEEE Proc.*

Treitel, S. and Lines, L.R., 1999. Past, present and future of geophysical inversion-a Y2K analysis. CREWES Research Report.

Turhan, M., 1995. Neural networks and computation of neural network weights and biases generalized delta rule and backpropagation of errors. http://www.discoby.com/pdf/neural_network.

Uchida, T., 1993. Smooth 2-D inversion for magnetotelluric data based on statistical criterion ABIC. *Journal of Geomagnetism and Geoelectricity*, **45**, 841-858.

van der Bean, M. and Jutten, C., 2000. Neural networks in geophysical applications. *Geophysics*, **65**, 1032-1047.

Vozoff, K., 1991. The magnetotelluric method. In: Nabighian, M.N. (Ed), Electromagnetic Method In Applied Geophysics. Society of Exploration Geophysicist, 641-712.

Wang, L.X. and Mendal, J.M., 1992. Adaptive minimum prediction error deconvolution and source wavelete estimation using Hopfield neural networks. *Geophysics*, **57**, 670-679.

Wannamaker, P.E., Stodt, J.A. and Rijo, L., 1987. A stable finite element solution for two-dimensional magnetotelluric modeling. *Geophysical Journal of the Royal Astronomical Society*, **88**, 277-296.

Wessel, P. and Smith, W.H.F., 2004. The Generic Mapping Tools (GMT). <http://www.gmt.soest.hawaii.edu>.

Zhdanov, M.N. and Keller, G.V., 1998. The Geoelectrical Methods in Geophysical Exploration, 873.

پیوست (الف)

پیوست (الف):

بخش اول - محاسبه معادلات موج الکترومغناطیسی برای هر قطبش

با در نظر گرفتن مدلی مطابق با شکل ۱-۲ و با فرض ثابت بودن تغییرات میدان در راستای x

(جهت امتداد) یعنی $\frac{\partial}{\partial x} = 0$ ، برای حالت TE میدان‌ها به صورت زیر می‌باشند [مرادزاده، ۱۹۹۸].

$$E = (E_x, 0, 0)$$

$$H = (0, H_y, H_z) \quad (\text{الف-۱})$$

$$E_x = E_x(y, z)$$

با استفاده از معادلات (۱-۲) و (۲-۲) معادلات زیر حاصل می‌شود:

$$\frac{\partial E_x}{\partial z} = -j\omega\mu H_y$$

$$\frac{\partial E_x}{\partial y} = j\omega\mu H_z \quad (\text{الف-۲})$$

$$E_x = \frac{1}{\sigma(y, z)} \left(\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} \right)$$

و نهایتاً معادلهٔ پخش برای این حالت به صورت زیر می‌باشد:

$$\nabla^2 E_x - j\omega\mu\sigma E_x = 0$$

$$\frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2} - j\omega\mu\sigma E_x = 0 \quad (\text{الف-۳})$$

به همین ترتیب برای حالت TM نیز داریم:

$$E = (0, E_y, E_z)$$

$$H_x = H_x(y, z) \quad (\text{الف-۴})$$

$$H = (H_x, 0, 0)$$

با استفاده از معادلات ماکسول (۱-۲) و (۲-۲) بدست می آید:

$$\frac{\partial H_x}{\partial z} = \sigma(y, z) E_y$$

$$\frac{\partial H_x}{\partial y} = -\sigma(y, z) E_z \quad (\text{الف-۵})$$

$$H_x = -\frac{1}{j\omega\mu} \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \right)$$

$$\frac{\partial^2 H_x}{\partial y^2} + \frac{\partial^2 H_x}{\partial z^2} - j\omega\mu\sigma H_x = 0$$

بخش دوم: رابطه بین امپدانس های تانسوری و امپدانس های اصلی

$$Z'_{xx} = \frac{1}{2} [(Z_{xx} + Z_{yy}) + (Z_{xy} + Z_{yx}) \sin 2\theta + (Z_{xx} - Z_{yy}) \cos 2\theta] \quad (\text{الف-۶})$$

$$Z'_{xy} = \frac{1}{2} [(Z_{xy} - Z_{yx}) + (Z_{xy} + Z_{yx}) \cos 2\theta - (Z_{xx} - Z_{yy}) \sin 2\theta] \quad (\text{الف-۷})$$

$$Z'_{yx} = \frac{1}{2} [(Z_{yx} - Z_{xy}) + (Z_{xy} + Z_{yx}) \cos 2\theta - (Z_{xx} - Z_{yy}) \sin 2\theta] = -Z'_{xy}(\theta \pm \pi/2) \quad (\text{الف-۸})$$

$$Z'_{yy} = \frac{1}{2} [(Z_{xx} + Z_{yy}) - (Z_{xy} + Z_{yx}) \sin 2\theta - (Z_{xx} - Z_{yy}) \cos 2\theta] = Z'_{xx}(\theta + \pi/2) \quad (\text{الف-۹})$$



Shahrood University of Technology
Faculty of Mining Engineering and Geophysics

Two-dimensional inverse modeling of magnetotelluric data using artificial neural networks

*This thesis submitted in part fulfillment of the degree of Master of Science in
Geophysics*

By:

F. Tahmassebi

Supervisors:

Dr. A. Moradzadeh

Dr. M. M. Fateh

Consultant:

M. Allamehzadeh

September 2004

Abstract

The magnetotelluric (MT) method is a natural source electromagnetic geophysical method that is used for petroleum, mineral resources and geothermal reservoirs exploration. This method is based on simultaneous measurement of the electrical and magnetic fields variations on the earth surface. To have a quantitative and meaningful interpretation, the MT data need to be modeled somehow. There are different inversion procedures that their precision depends on data quality, noise level and type of inversion methods. The common inversion require some parameter values to be found in advance from the other geophysical method or geological information. As the governing equations of MT field are very complex, hence all above methods often use numerical methods to solve equations. Once the dimension of model and its complexity are increased the accuracy of the calculation is reduced and more time is required to solve the problems so that the process become a time consuming task.

With attention to the Artificial Neural Networks (ANN) properties and their ability to solve non-linear problems and function approximation, it seems they could be a possible new solution for these difficulties.

In this study a multilayer perceptron (MLP) with back propagation (BP) learning algorithm was used to inverse MT data. In order to train, test, and validate the network, 498 sets of training samples have synthetically been created by a computer based MT data modeling software named EMIXMT2D. Synthetic data, which is really the response of the 498 two-dimensional models, are apparent resistivity and impedance phase of 11 frequencies at 9 stations for two polarization modes (i.e. Transverse Electric (TE) and Transverse Magnetic (TM) polarization modes). Once the required networks have been designed and tested, the effect of various input data (xy and yx apparent resistivity and phase) has been investigated using 5 different networks. To simulate the field data situation, a series of randomly distributed noise has been included on data and their effects have been studied on inverse modeling of data.

The results indicate that if ANN are designed and trained properly, then they would be capable enough to perform two-dimensional inverse modeling of MT data. The various evaluation methods in this research work show that once the designed networks have been trained properly they are able to perform the inverse modeling precisely in a short time, which is less than a few minutes. At the end, the performance of one of the designed networks (MT net) has been evaluated by a set of field MT data and its results has been compared with those produced by a common smooth inverse (Rapid Relaxation Inversion, RRI) method. The results obtained from two different procedures are in close agreement, whereas, the extension of the anomalous zone are better determined by ANN inversion of data.