

FONTES DE ERRO

Os erros que afectam as observações GPS podem ter origem nos satélites, nas antenas e receptores e no meio de propagação. De um modo geral, os erros podem ser modelados eficazmente, mas há casos de erros que não podem ser modelados com facilidade.

Os erros com origem nos satélites incluem os erros orbitais (erros das efemérides dos satélites), erros dos relógios dos satélites e o acesso selectivo.

Os erros com origem nos receptores incluem o erro do relógio (muito maior do que o erro do relógio do satélite) e o ruído do receptor (circuitos electrónicos). Os erros de multitrajecto e variação do centro de fase da antena estão associados à antena.

Finalmente, há que considerar os erros associados à atmosfera (ionosfera e troposfera), que influencia a propagação dos sinais rádio.

Erros dependentes dos satélites

Erros das órbitas

Na determinação de órbitas, a imprecisão das posições e velocidades iniciais dos satélites (necessárias à resolução das equações diferenciais que descrevem o movimento de um satélite) - que é causada por erros nas coordenadas das estações de rastreio e a utilização de modelos de força imprecisos (na modelação das perturbações da órbita do satélite) -, resultam em erros das efemérides dos satélites GPS. De entre os modelos de força, aquele relativo à modelação da pressão de radiação solar é o que mais contribui para o erro final.

O efeito dos erros nas órbitas manifesta-se de forma diferente no posicionamento absoluto e no posicionamento relativo. No caso do posicionamento relativo, o erro resultante numa base pode ser determinado a partir de uma relação simples:

$$\frac{dB}{B} = \frac{dr}{\rho}, \quad (57)$$

onde dB é o erro no vector-base, B , e dr é o erro no vector distância satélite-receptor, ρ . Esta expressão dá origem a uma estimacão pessimista, mas apresenta a vantagem da simplicidade relativamente a outras “regras de ouro”.

Como exemplo, e para o caso de um satélite GPS ($\rho \approx 20\,000$ km), um erro de 2 m na órbita produz um erro de 1 cm numa base de 100 km.

Para além das órbitas radiodifundidas, o utilizador poderá utilizar órbitas pós-processadas. As órbitas radiodifundidas apresentam uma incerteza da ordem de 1 m, enquanto que a incerteza associada às órbitas precisas, tal como aquelas que são geradas pelo IGS, é actualmente de cerca de 2.5 cm (órbitas finais IGS).

Erros dos relógios

Embora os relógios dos satélites sejam muito precisos, não são de modo algum perfeitos. No entanto, o erro do relógio do satélite poderá ser eliminado através da formação de diferenças duplas, dado que o erro é comum aos receptores que observam esse satélite. O mesmo efeito é conseguido incluindo parâmetros que modelem o comportamento do relógio em cada época no processo de ajustamento, utilizando geralmente um processo estocástico. Finalmente, por aplicação da correcção ao relógio constante na mensagem de navegação, pode-se reduzir o erro no relógio do satélite a alguns nanosegundos. O IGS também produz informação sobre os relógios dos satélites (com incertezas de ~ 20 ps para o produto final).

Acesso selectivo

O efeito do acesso selectivo deve ser incluído neste grupo de erros. Tal como foi referido anteriormente, a técnica δ altera a informação do relógio do satélite e a técnica ϵ provoca alterações na informação orbital. O acesso selectivo prejudica basicamente o posicionamento absoluto, sendo essencialmente eliminado com a utilização de técnicas diferenciais.

Erros dependentes da antena-receptor

Multitrajecto

O erro de multitrajecto resulta da interferência mútua de dois ou mais sinais, emitidos por uma fonte comum, que percorrem trajectos de diferentes comprimentos, devido a reflexões. Os sinais que desta forma chegam ao receptor vêm desfasados e interferem uns com os outros. Para além das superfícies reflectoras próximas (edifícios, carros, etc.), os outros factores que afectam o multitrajecto são a geometria dos satélites e as características das antenas. Os efeitos do multitrajecto na pseudo-distância são geralmente da ordem dos poucos centímetros, mas poderá atingir valores mais significativos (1-2 m, para o código C/A); para a fase, a distorção do sinal original pode ter um valor máximo de um quarto de ciclo.

O erro de multitrajecto pode ser detectado por análise da correlação elevada existente nos resíduos do ajustamento em dois dias consecutivos, devido à repetição, de dia sideral para dia sideral, da geometria satélite-antena-reflector. Infelizmente, a modelação do efeito do multitrajecto não é fácil. Embora as antenas actuais sejam menos sensíveis ao efeito do multitrajecto, o problema do multitrajecto continua a constituir um factor limitativo na precisão do GPS. Em modo relativo estático, o erro de multitrajecto pode ser reduzido usando sessões de observação longas. De acordo com Genrich and Bock [1992], um período de ocupação de 30 minutos é geralmente suficiente para reduzir o erro de multitrajecto a valores inferiores a 1 mm. Para posicionamento cinemático e rápido-estático (ver “Métodos de Posicionamento”), o erro de multitrajecto pode ser significativo. Existem actualmente uma grande variedade de antenas (geralmente utilizadas nas aplicações GPS à geodesia de precisão e geodinâmica) que proporcionam um baixo nível de multitrajecto (antenas *choke ring* – ver Figura 15).

Variações do centro de fase da antena

As variações do centro de fase da antena (PCV) são função da construção e características da antena e dependem do ângulo de direcção do sinal observado. Estas variações podem atingir alguns centímetros para alguns tipos de antenas. Quando utilizados em modo de rede (ou base simples), as antenas deverão ser orientadas espacialmente de modo semelhante (para Norte), de forma a que se minimize este efeito. As antenas *choke ring* já mencionadas possuem um baixo grau de variação do centro de fase.



Figura 15 – Antena do tipo choke ring.

A determinação das variações do centro de fase de uma antena pode ser feita usando calibração, podendo esta ser feita em modo relativo ou em modo absoluto.

Na **calibração relativa** em campo, montam-se duas antenas (uma de referência e outra a calibrar) em dois pontos próximos, com coordenadas conhecidas. Assume-se que a antena de referência tem PCV nulas. O processamento das observações e inspeção dos resíduos permite estimar as PCV da antena a calibrar.

A **calibração absoluta** pode ser feita num laboratório especialmente desenhado (câmara anecóica) ou através de robotização.

Na calibração em câmara anecóica, usa-se um sinal GPS artificial e faz-se deslocar a antena GPS em diferentes posições. É uma técnica pouco prática, devido ao reduzido número de câmaras anecóicas.

Na calibração com robotização (Figura 16), há um mecanismo robotizado que faz girar a antena em todas as elevações e azimutes. O processamento das observações em dois dias siderais diferentes permite fazer a determinação das PCV. É um processo utilizado actualmente. As PCV são distribuídas à comunidade científica utilizando um formato padrão, denominado ANTEX (*Antenna Exchange Format*).



Figura 16 - Calibração de antenas GPS em modo absoluto.

Ruído do receptor

O conhecimento do desvio padrão associado a cada medição, geralmente conhecido por ruído do receptor, é necessário na estimação das precisões esperadas no posicionamento relativo com GPS. O ruído do receptor associado a cada tipo de medição resulta das limitações inerentes à electrónica dos componentes do receptor. Como regra geral, pode-se dizer que o ruído das observações é cerca de 1% do comprimento de onda do sinal, ou seja, cerca de 2 mm para as medições de fase. Os receptores mais recentes apresentam ruído bastante inferior a este valor indicativo (cerca de uma ordem de magnitude).

Assumindo medições de fase não correlacionadas, e assumindo 2 mm como ruído na medição de fase, este valor será propagado para 3 mm, na formação de diferenças simples, e para 4 mm, na formação de diferenças duplas. Usando a combinação linear da ionosfera livre em diferenças duplas, o ruído esperado será de cerca de 12 mm.

Erro do relógio do receptor

A modelação do erro do relógio do receptor pode ser feito de modo idêntico à usada no erro do relógio do satélite, quer estimando parâmetros de relógio (desfasamento e deriva) quer formando diferenças de observações, que eliminem o seu efeito.

Embora não seja um problema inerente quer à antena quer ao receptor, é de referir o erro associado à medição da altura da antena relativamente ao ponto-estação. Este erro é relativamente pequeno, quando comparado com as variações do centro de fase. Medições cuidadas deverão proporcionar uma precisão da ordem de 1 a 3 mm.

Erros dependentes do meio de propagação

Ionosfera

A ionosfera é a parte da atmosfera terrestre onde os iões e electrões, resultantes da radiação ultravioleta emitida pelo Sol, estão presentes em quantidade suficiente para afectar a propagação das ondas rádio. Está compreendida entre os 50 km e os 1000 km e afecta os sinais GPS, provocando atrasos de propagação do sinal que podem atingir várias dezenas de metros, para observações próximas do horizonte e em períodos de maior actividade solar. Os limites mencionados são meramente indicativos, dado o carácter variável da ionosfera. Embora seja habitual dizer que o limite superior da ionosfera se situa a cerca de 1000 km de altitude, não existe nenhuma fronteira real entre a ionosfera e as camadas exteriores do campo magnético da Terra.

Caracterização da Ionosfera

A variabilidade da ionização da atmosfera terrestre conduz a uma divisão da ionosfera em diferentes camadas, com limites não muito bem definidos (são indicados valores meramente indicativos).

Camada D – Varia entre os 50 km e os 90 km. A principal fonte de ionização é a radiação cósmica (radiação formada por partículas elementares de origem extraterrestre, resultantes de interacções nucleares entre partículas altamente energéticas emitidas pelo Sol e outras estrelas e as moléculas de gás da atmosfera). A camada D é composta de iões carregados negativa e positivamente e electrões livres. Durante a noite, os electrões desta camada podem combinar-se com átomos e moléculas dando origem a iões negativos, o que faz com que a camada D desapareça. Durante o dia, os electrões tendem a separar-se de novo, devido à radiação solar, fazendo com que a camada D reapareça. É caracterizada por uma baixa densidade de electrões, apresentando como valores típicos $6.1 \times 10^8 - 13.1 \times 10^8$ electrões/m³, para as regiões de latitude média e ao meio dia local, estando a variação associada à actividade solar [Komjathy, 1997]. Em condições normais, o grau de ionização não é suficientemente forte para afectar a propagação de ondas rádio.

Camada E – Varia entre os 90 km e os 140 km. A ionização na camada E apresenta uma grande correlação com a radiação solar directa, tornando-a quase inexistente durante a noite, uma vez que os electrões voltam a combinar-se com os iões positivos. Durante o dia, a densidade de electrões é cerca de uma ordem de magnitude superior à observada na camada D (1.3×10^{11} – 1.7×10^{11} electrões/m³, para as condições consideradas anteriormente [Komjathy, 1997]). Por vezes, e dependendo das condições de vento solar e energia absorvida durante o dia, a camada E pode permanecer esporadicamente à noite, originando a **camada E-Esporádica** (Es). Esta camada fina é altamente variável e contém uma concentração de electrões anormalmente elevada.

Camada F – A camada F tem como limite inferior 140 km e compreende duas regiões de características diferentes: uma região mais baixa, denominada F1 – que apenas existe durante o dia – e uma camada superior, denominada F2, caracterizada por uma grande concentração de electrões, que atinge o máximo na região compreendida entre os 250 km e os 500 km.

A principal fonte de ionização da camada F1 é a radiação ultravioleta do Sol, pelo que é observada apenas durante o dia. A densidade de electrões nesta camada, para latitudes médias e ao meio-dia local, é de 2.3×10^{11} – 3.3×10^{11} electrões/m³. Para as mesmas condições, a densidade de electrões para a camada F2 é de 2.8×10^{11} – 5.2×10^{11} electrões/m³ [Komjathy, 1997]. Embora não desapareça totalmente durante a noite, a densidade de electrões durante este período decresce cerca de 1 ordem de magnitude.

A altitude da camada F varia em função da hora do dia, da época do ano e é muito sensível ao ciclo de manchas solares (“*sunspot*”). É também muito sensível às perturbações ionosféricas rápidas devidas a fenómenos como as chamas solares e as perturbações magnéticas. A camada F é uma região de grande interesse para comunicações rádio, uma vez que constitui um meio de reflexão para altas frequências, permitindo a sua propagação a longas distâncias.

Ionosfera Superior (“*Topside*”) – Esta camada tem como limite inferior a altitude correspondente à máxima densidade de electrões da camada F2 e estende-se até uma altitude de transição (onde os iões O⁺ são menos numerosos que os iões H⁺ e He⁺), que pode chegar aos 1100 km. Acima desta zona de transição, a ionização é tão fraca que tem pouca influência na propagação das ondas rádio. Por analogia, a camada abaixo da Ionosfera Superior é a **Ionosfera Inferior (“*Bottomside*”)**.

A ionosfera pode ainda sofrer perturbações bruscas, devido a fenómenos como as chamadas solares e perturbações magnéticas. Para determinados períodos do dia (principalmente após o pôr-do-sol), algumas regiões da ionosfera (principalmente a região F das altas e baixas latitudes) podem tornar-se bastante turbulentas, ou seja, podem-se formar estruturas de pequena escala (variação de centímetros a metros) na ionosfera ambiente (com escalas da ordem das dezenas de quilómetros) que podem afectar significativamente a propagação das ondas rádio. A este fenómeno dá-se o nome de **cintilação ionosférica** e tem como consequência uma rápida flutuação da fase e da intensidade de um sinal de rádio.

Principais regiões geográficas da ionosfera

Existem três grandes regiões ionosféricas: alta latitude, latitude média e equatorial.

Região equatorial – É caracterizada por ter os valores mais elevados de densidade de electrões, com efeitos pronunciados de amplitude e cintilação de fase. O efeito combinado do alto nível de radiação solar e dos campos eléctrico e magnético da Terra (o facto do campo geomagnético fluir horizontalmente sobre o Equador magnético gera uma condutividade eléctrica anormalmente elevada sobre o Equador) origina uma subida de electrões que se movem ao longo das linhas horizontais do campo magnético, um fenómeno conhecido por “efeito de fonte” (“*fountain effect*”), que distorce a forma geral da ionosfera. Os electrões movem-se até próximo das latitudes geomagnéticas dos $10^\circ - 20^\circ$, provocando uma grande concentração de electrões, denominadas *anomalias equatoriais*.

Latitude média – É a menos variável e a menos perturbada das diferentes regiões ionosféricas, não sofrendo o efeito do campo magnético horizontal próprio da região equatorial.

Alta latitude – É caracterizada por ter ionização por colisão, para além da fotoionização, devido ao facto das linhas do campo geomagnético serem aproximadamente verticais nesta região, provocando uma descida das partículas carregadas até à camada E (~100 km). Estas partículas podem colidir com gases atmosféricos electricamente neutros, provocando um aumento de concentração de electrões, um fenómeno associado à actividade das auroras (as auroras resultam de interacção entre a magnetosfera, a ionosfera e a atmosfera).

Refracção Ionosférica

De acordo com o princípio de Fermat, o trajecto percorrido por uma onda electromagnética entre dois pontos arbitrários (L) é dado por:

$$L = \int_{\text{raio}} n ds, \quad (58)$$

onde ds é o elemento diferencial de comprimento ao longo do trajecto e n é o índice de refração do meio. Este trajecto corresponde ao trajecto mínimo sob o ponto de vista de propagação do sinal electromagnético (**trajecto electromagnético**), mas não corresponde ao trajecto mínimo em termos de comprimento. Dado que a trajectória de um sinal na atmosfera terrestre atravessa camadas com índices de refração diferentes, o que obriga a mudança de direcção do raio, o trajecto efectivo é superior àquele que seria obtido se percorresse um percurso rectilíneo, denominado **trajecto geométrico**, G :

$$G = \int_{\text{vac}} ds, \quad (59)$$

sendo este integral calculado no vácuo. A diferença entre os comprimentos dos trajectos electromagnético e geométrico ionosféricos constitui a **refracção ionosférica**:

$$d_{\text{ion}} = \int_{\text{raio}} n ds - \int_{\text{vac}} ds. \quad (60)$$

Se designarmos por S o **comprimento geométrico** do percurso efectivo:

$$S = \int_{\text{raio}} ds, \quad (61)$$

ao somar e subtrair à Equação (60), obtemos:

$$d_{\text{ion}} = \int_{\text{raio}} n ds - \int_{\text{raio}} ds + \int_{\text{raio}} ds - \int_{\text{vac}} ds, \quad (62)$$

ou

$$d_{\text{ion}} = \int_{\text{raio}} (n-1) ds + \left[\int_{\text{raio}} ds - \int_{\text{vac}} ds \right]. \quad (63)$$

Nesta equação, o primeiro termo do lado direito representa o excesso de percurso provocado pelo atraso de propagação do sinal, denominado **atraso ionosférico**, enquanto que o segundo termo é o excesso de percurso provocado pela curvatura do raio, denominado **atraso geométrico** ou **curvatura do raio** (*ray bending*). O cálculo do atraso ionosférico pressupõe, desta forma, que se conheça o índice de refração ao longo do trajecto.

Atraso ionosférico

O índice de refração de um meio é dado por:

$$n = \frac{c}{v}, \quad (64)$$

onde c é a velocidade de propagação no vácuo e v a velocidade de propagação no meio. Numa situação ideal de propagação das ondas electromagnéticas no vácuo, o índice de refração seria igual a 1 e as velocidades de propagação de fase e de grupo seriam iguais. Como parte do trajecto de um sinal rádio é feito através da atmosfera, as velocidades de propagação para a fase e para o grupo, v_ϕ e v_g , são função dos índices de refração da fase, n_ϕ , e do grupo, n_g :

$$v_\phi = \frac{c}{n_\phi} \quad (65)$$

$$v_g = \frac{c}{n_g}. \quad (66)$$

Uma onda pura é afectada pelo índice de refração de fase do meio: velocidade de fase. Esta velocidade é dada pela relação:

$$v_\phi = \lambda f. \quad (67)$$

Um sinal modulado é afectado pelo índice de refração de grupo (grupo de ondas com frequências ligeiramente diferentes superimpostas numa mesma onda portadora): velocidade de grupo. A velocidade de grupo traduz a velocidade à qual a informação é transmitida e é dada por:

$$v_g = -\frac{df}{d\lambda} \lambda^2. \quad (68)$$

Que relação existe entre a velocidade de fase e a velocidade de grupo?

Começemos por calcular o diferencial para a velocidade de fase. Dado que

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i, \quad (69)$$

tem-se

$$dv_{\phi} = fd\lambda + \lambda df. \quad (70)$$

Dividindo por $d\lambda$ e reorganizando, tem-se:

$$\frac{df}{d\lambda} = \frac{1}{\lambda} \frac{dv_{\phi}}{d\lambda} - \frac{f}{\lambda}. \quad (71)$$

Substituindo na Equação (68), tem-se:

$$v_g = -\lambda \frac{dv_{\phi}}{d\lambda} + f\lambda. \quad (72)$$

Introduzindo a Equação (67), tem-se:

$$v_g = v_{\phi} - \lambda \frac{dv_{\phi}}{d\lambda}, \quad (73)$$

que é conhecida como **Equação de Rayleigh** e que estabelece a relação entre a velocidade de grupo e a velocidade de fase.

Como alternativa, e dado que

$$\lambda = \frac{c}{f}, \quad (74)$$

tem-se,

$$d\lambda = -\frac{\lambda}{f} df \quad (75)$$

ou, de forma equivalente,

$$\frac{d\lambda}{\lambda} = -\frac{df}{f} \quad (76)$$

Introduzindo a Equação (76) na Equação (73), obtemos:

$$\mathbf{v}_g = \mathbf{v}_\phi + f \frac{d\mathbf{v}_\phi}{df}. \quad (77)$$

A partir da Equação (65), tem-se:

$$d\mathbf{v}_\phi = \frac{\partial \mathbf{v}_\phi}{\partial \mathbf{n}_\phi} d\mathbf{n}_\phi = -\frac{c}{\mathbf{n}_\phi^2} d\mathbf{n}_\phi, \quad (78)$$

ou ainda, dividindo por $d\lambda$,

$$\frac{d\mathbf{v}_\phi}{d\lambda} = -\frac{c}{\mathbf{n}_\phi^2} \frac{d\mathbf{n}_\phi}{d\lambda}. \quad (79)$$

De forma semelhante, tem-se:

$$\frac{d\mathbf{v}_\phi}{df} = -\frac{c}{\mathbf{n}_\phi^2} \frac{d\mathbf{n}_\phi}{df}. \quad (80)$$

Substituindo a Equação (80) na Equação (73), e atendendo às Equações (65) e (66), tem-se:

$$\frac{c}{\mathbf{n}_g} = \frac{c}{\mathbf{n}_\phi} + \lambda \frac{d\mathbf{n}_\phi}{d\lambda}. \quad (81)$$

Dividindo ambos os termos por c e reorganizando, tem-se:

$$\frac{1}{\mathbf{n}_g} = \frac{1}{\mathbf{n}_\phi} \left(1 + \lambda \frac{1}{\mathbf{n}_\phi} \frac{d\mathbf{n}_\phi}{d\lambda} \right). \quad (82)$$

Utilizando a relação matemática seguinte:

$$(1-x)^{-1} = 1 + x + x^2 + \dots \quad |x| < 1, \quad (83)$$

a Equação (82) pode ser simplificada (retendo apenas os dois primeiros termos da expansão) e apresentada na seguinte forma:

$$\mathbf{n}_g = \mathbf{n}_\phi - \lambda \frac{d\mathbf{n}_\phi}{d\lambda}, \quad (84)$$

que expressa a relação que existe entre o índice de refração de grupo e o índice de refração de fase.

Em alternativa, tem-se:

$$n_g = n_\phi + f \frac{dn_\phi}{df}. \quad (85)$$

O índice de refração para a fase pode ser aproximado pela seguinte expansão em série:

$$n_\phi = 1 + \frac{c_1}{f^2} + \frac{c_2}{f^3} + \frac{c_3}{f^4} + \dots, \quad (86)$$

onde os coeficientes c_i ($i = 1, \dots, n$) são apenas função do número de electrões por m^3 (N_e - densidade de electrões) ao longo do trajecto de propagação do sinal. Retendo apenas os dois primeiros termos da série, tem-se:

$$n_\phi = 1 + \frac{c_1}{f^2}, \quad (87)$$

Diferenciando, vem:

$$dn_\phi = \frac{\partial n_\phi}{\partial f} df = -\frac{2c_1}{f^3} df,$$

ou seja,

$$\frac{dn_\phi}{df} = -\frac{2c_1}{f^3}. \quad (88)$$

Substituindo na Equação (85) e simplificando, obtemos:

$$n_g = 1 - \frac{c_1}{f^2}, \quad (89)$$

ou, em alternativa,

$$n_\phi = 1 + \frac{c_1}{f^2}. \quad (90)$$

O valor de c_1 (em Hz^2) é de $-40.28 N_e$. Assim, o índice de refração da fase é dado, em primeira aproximação, pela chamada **fórmula de Appleton-Hartree**:

$$n_\phi = 1 - 40.28 \frac{N_e}{f^2}. \quad (91)$$

Para o índice de refração de grupo, tem-se:

$$n_g = 1 + 40.28 \frac{N_e}{f^2}. \quad (92)$$

Podemos concluir que a ionosfera aumenta a velocidade de propagação da fase portadora para além da velocidade da luz, **avanço de fase**, e reduz de igual quantidade a velocidade de propagação da onda modulada, **atraso de grupo**.

Usando as Equações (90) e (91), e calculando o atraso ionosférico apenas segundo a direcção do zénite (elemento diferencial de trajecto dz), obtemos respectivamente:

$$d_{\text{ion}_g}^z = \frac{40.28}{f^2} \int_{\text{raio}} N dz \quad (93)$$

$$d_{\text{ion}_\phi}^z = -\frac{40.28}{f^2} \int_{\text{raio}} N dz \quad (94)$$

O integral $\int_{\text{raio}} N dz$ é designado como **conteúdo total de electrões** (TEC -

Total Electron Content) e corresponde ao número de electrões numa coluna de área de secção transversal de 1 m^2 , ao longo do trajecto receptor-satélite. Em geral, o TEC tem unidades 10^{16} electrões/ m^2 (valor correspondente a uma unidade TEC, denominada **TECU**). O TEC é uma quantidade calculada na direcção do zénite; para relacionar este valor para um ângulo de elevação qualquer é necessário introduzir uma função de mapeamento ou factor de obliquidade. O atraso ionosférico depende assim do TEC, pelo que é função da localização geográfica, hora do dia e actividade solar, entre outros parâmetros.

Assim, o atraso ionosférico para a fase pode ser escrito sob a forma:

$$d_{\text{ion}_\phi}^z = -40.28 \frac{\text{TEC}}{f^2}. \quad (95)$$

Procedendo de forma idêntica para o atraso ionosférico de grupo, vem:

$$d_{\text{ion}_g}^z = 40.28 \frac{\text{TEC}}{f^2}. \quad (96)$$

Em resumo, os atrasos ionosféricos de grupo e fase são iguais e de magnitude oposta, e ambos dependentes da frequência do sinal rádio (meio dispersivo). A simples combinação de observações de pseudo-distância e fase feitas na mesma frequência podem, desta forma, ser utilizadas para remover os efeitos ionosféricos (embora a observação resultante tenha um ruído muito elevado).

Determinação do atraso ionosférico a partir de duas frequências

A propriedade de dispersão da ionosfera para sinais rádio permite eliminar o efeito de primeira ordem, fazendo a combinação linear de frequências, como já foi referido anteriormente. É também possível, partindo de observações em duas frequências, obter os valores do atraso ionosférico.

Consideremos então o efeito de primeira ordem do atraso ionosférico para L1 e L2 (índices 1 e 2, respectivamente):

$$d_{\text{ion}_1} = \frac{40.28}{f_1^2} \text{TEC}$$

e

$$d_{\text{ion}_2} = \frac{40.28}{f_2^2} \text{TEC},$$

onde f_1 e f_2 são, como habitualmente, as frequências associadas às portadoras L1 e L2, respectivamente.

A partir destas relações, tem-se:

$$d_{\text{ion}_1} f_1^2 = d_{\text{ion}_2} f_2^2, \quad (97)$$

ou seja,

$$\begin{cases} d_{\text{ion}_1} = \frac{d_{\text{ion}_2} f_2^2}{f_1^2} \\ d_{\text{ion}_2} = \frac{d_{\text{ion}_1} f_1^2}{f_2^2} \end{cases} \quad (98)$$

As equações de pseudo-distância para L1 e L2 são, respectivamente (ignorando outros erros, ε),

$$p_1 = \rho + c(dt - dT) + d_{\text{ion}_1} + d_{\text{trop}}$$

$$p_2 = \rho + c(dt - dT) + d_{\text{ion}_2} + d_{\text{trop}}$$

pelo que:

$$p_2 - p_1 = d_{\text{ion}_2} - d_{\text{ion}_1}. \quad (99)$$

Utilizando a Equação (98), tem-se

$$p_2 - p_1 = \frac{d_{\text{ion}_1} f_1^2}{f_2^2} - d_{\text{ion}_1} \quad (100)$$

$$p_2 - p_1 = d_{\text{ion}_1} \left(\frac{f_1^2 - f_2^2}{f_2^2} \right), \quad (101)$$

ou seja,

$$d_{\text{ion}_1} = \left(\frac{f_2^2}{f_1^2 - f_2^2} \right) (p_2 - p_1). \quad (102)$$

Da forma análoga, obtém-se:

$$d_{\text{ion}_2} = \left(\frac{f_1^2}{f_2^2 - f_1^2} \right) (p_2 - p_1). \quad (103)$$

No caso das observações de fase, tem-se (ignorando outros erros, ε):

$$\Phi_1 = \rho + c(dt - dT) + \lambda_1 N_1 - d_{\text{ion}_1} + d_{\text{trop}}$$

$$\Phi_2 = \rho + c(dt - dT) + \lambda_2 N_2 - d_{\text{ion}_2} + d_{\text{trop}}$$

$$\Phi_2 - \Phi_1 = \lambda_2 N_2 - \lambda_1 N_1 - d_{\text{ion}_2} + d_{\text{ion}_1} \quad (104)$$

Utilizando a Equação (98), tem-se

$$\Phi_2 - \Phi_1 = \lambda_2 N_2 - \lambda_1 N_1 - \frac{d_{\text{ion}_1} f_1^2}{f_2^2} + d_{\text{ion}_1} \quad (105)$$

$$(\Phi_2 - \Phi_1) - (\lambda_2 N_2 - \lambda_1 N_1) = d_{\text{ion}_1} \left(1 - \frac{f_1^2}{f_2^2} \right) = d_{\text{ion}_1} \left(\frac{f_2^2 - f_1^2}{f_2^2} \right)$$

$$d_{\text{ion}_1} = \frac{f_2^2}{f_2^2 - f_1^2} [(\Phi_2 - \Phi_1) - (\lambda_2 N_2 - \lambda_1 N_1)] \quad (106)$$

$$d_{\text{ion}_2} = \frac{f_1^2}{f_2^2 - f_1^2} [(\Phi_2 - \Phi_1) - (\lambda_2 N_2 - \lambda_1 N_1)]. \quad (107)$$

Devido ao facto dos valores de N_1 e N_2 serem desconhecidos, as equações de fase não permitem a determinação do valor absoluto dos atrasos ionosféricos, contrariamente às equações de pseudo-distância. No entanto, se não houver saltos de ciclo, as observações de fase permitem avaliar as variações relativas no atraso ionosférico. Partindo destas características, podemos combinar os atrasos ionosféricos obtidos com pseudo-distâncias com aqueles obtidos a partir de fases, tirando partido do carácter absoluto dos primeiros (embora com grande dispersão, devido ao ruído das observações de pseudo-distância) e a boa qualidade dos segundos. Para tal, basta determinar a diferença média entre os dois atrasos e somar essa média ao atraso ionosférico determinado a partir das fases, gerando um **atraso de grupo suavizado com fase**. Dada a relação directa, a conhecimento do atraso ionosférico permite o cálculo do TEC. Os mapas globais de TEC são outro dos produtos obtidos do GPS, sendo distribuídos à comunidade científica sob um formato padrão, denominado IONEX (*Ionospheric Map Exchange Format*). A Figura 17 mostra um desses mapas, produzido pelo JPL.

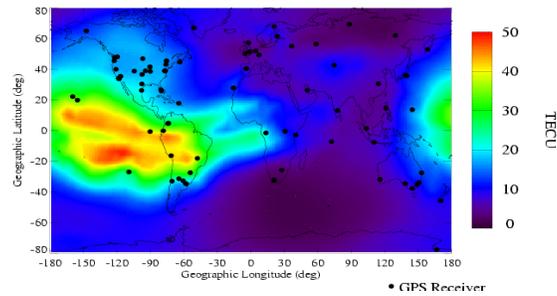


Figura 17 - Mapa global de TEC para o dia 2 de Novembro de 2005.

Modelos ionosféricos

Para os utilizadores de frequência única a modelação da ionosfera é mais complexa. Têm sido desenvolvidos vários modelos para modelar o efeito da ionosfera (teóricos e empíricos). A mensagem de navegação GPS contém coeficientes para serem usados num modelo ionosférico desenvolvido por Klobuchar (modelo ionosférico radiodifundido ou **modelo de Klobuchar**). Este modelo permite remover cerca de 50% (emq) do efeito global da ionosfera. O algoritmo correspondente ao modelo de Klobuchar [ICD, 1993] está representado na Figura 19. As variáveis envolvidas neste modelo constam da Tabela 1. A tabela dá os valores para o atraso ionosférico em L1 (a determinação para L2 é imediata, tendo em consideração a relação de frequências – ver Equação (98)). Na mesma tabela é feita referência ao ponto ionosférico e sub-ionosférico. O **ponto ionosférico** ou **ponto de incisão ionosférico** (IPP - *Ionospheric Pierce Point*) é definido pela intersecção do vector satélite-receptor e uma camada de altitude média (350 km, no caso do modelo de Klobuchar). O **ponto sub-ionosférico** é a projecção do ponto ionosférico na superfície da Terra (ver Figura 18).

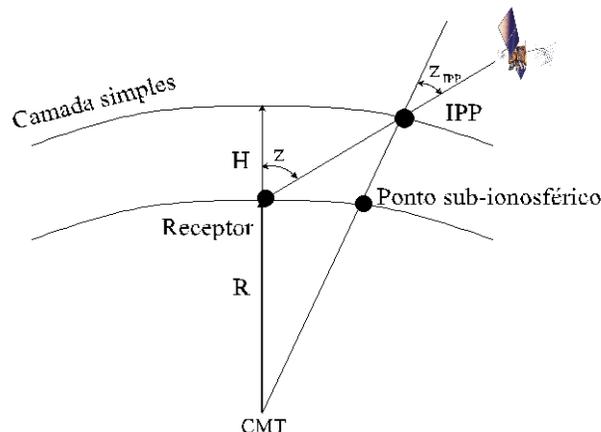


Figura 18 - O ponto sub-ionosférico e o IPP.

$$\begin{aligned}
 F &= 1 + 16.0 \times (0.53 - \varepsilon)^3 \\
 \Psi &= \frac{0.0137}{(\varepsilon + 0.11)} - 0.022 \\
 \begin{cases} \varphi_i = \varphi + \Psi \cos A & \text{se } |\varphi_i| \leq 0.416 \\ \varphi_i = 0.416 & \text{se } \varphi_i > 0.416 \\ \varphi_i = -0.416 & \text{se } \varphi_i < -0.416 \end{cases} \\
 \lambda_i &= \lambda + \Psi \frac{\sin A}{\cos \varphi_i} \\
 \phi_i &= \varphi_i + 0.064 \cos(\lambda_i - 1.167) \\
 \begin{cases} t = 43200\lambda_i + T & \text{se } 0 \leq t \leq 86400 \text{ s} \\ t = 43200\lambda_i + T + 86400 \text{ s} & \text{se } t < 0 \text{ s} \\ t = 43200\lambda_i + T - 86400 \text{ s} & \text{se } t > 86400 \text{ s} \end{cases} \\
 \begin{cases} P = \sum_{i=1}^4 \beta_{i-1} \phi^{i-1} & \text{se } P \geq 72000 \text{ s} \\ P = 72000 \text{ s} & \text{se } P < 72000 \text{ s} \end{cases} \\
 x &= \frac{2\pi(t - 50400)}{P} \\
 \begin{cases} Q = \sum_{i=1}^4 \alpha_{i-1} \phi^{i-1} & \text{se } Q \geq 0 \\ Q = 0 & \text{se } Q < 0 \end{cases} \\
 \begin{cases} d_{\text{ion}_{1g}} = 5 \times 10^{-9} \cdot c \cdot F & \text{se } |x| \geq 1.57 \\ d_{\text{ion}_{1g}} = c \cdot F \cdot \left(5 \times 10^{-9} + Q \cdot \left(1 - \frac{x^2}{2} + \frac{x^4}{24} \right) \right) & \text{se } |x| < 1.57 \end{cases}
 \end{aligned}$$

Figura 19 – Modelo ionosférico de Klobuchar.

Tabela 1 - Variáveis utilizadas no modelo de Klobuchar.

	UNIDADES	TIPO
Constantes		
c	ms ⁻¹	velocidade da luz no vácuo (299792458 ms ⁻¹)
Parâmetros transmitidos na mensagem de navegação		
α_i	–	Coefficientes de um polinómio do 3º grau que representa a amplitude do atraso ionosférico vertical (página 18 da subestrutura 4 da mensagem de navegação)
β_i	–	Coefficientes de um polinómio do 3º grau que representa o período do modelo (página 18 da subestrutura 4 da mensagem de navegação)
Parâmetros gerados pelo utilizador		
ε	semi-círculos	ângulo de elevação entre o utilizador e o satélite
A	semi-círculos	ângulo azimutal entre o utilizador e o satélite (azimute), contado positivamente no sentido dos ponteiros do relógio, a partir do Norte (verdadeiro)
φ	semi-círculos	latitude geodésica do utilizador
λ	semi-círculos	longitude geodésica do utilizador
T	s	aproximação ao tempo GPS, dado pelo receptor
Parâmetros calculados		
t	s	tempo local no ponto sub-ionosférico
φ_i	semi-círculos	latitude geodésica do ponto sub-ionosférico
λ_i	semi-círculos	longitude geodésica do ponto sub-ionosférico
ϕ	semi-círculos	latitude geomagnética do ponto sub-ionosférico
Ψ	semi-círculos	ângulo geocêntrico entre a estação GPS e o ponto sub-ionosférico
F	–	factor de obliquidade (função de mapeamento)
x	radianos	fase
P	s	período
Q	s	amplitude

Funções de mapeamento

A determinação do atraso ionosférico para um ângulo de elevação (ou ângulo zenital) genérico requer, como já foi referido, a introdução de uma **função de mapeamento**, também conhecida como **factor de obliquidade**. Um exemplo de uma dessas funções é aquela que está associada ao modelo de Klobuchar. A função de mapeamento mais simples é aquela que resulta do pressuposto que a Terra é plana e a ionosfera é homogénea e concentrada numa fina camada. Neste caso, a função de mapeamento é simplesmente a secante do ângulo zenital, calculado no ponto de incidência ionosférico (z_{ipp}):

$$F = \frac{1}{\cos z_{ipp}}. \quad (108)$$

Este ângulo zenital pode relacionar-se com o ângulo zenital do satélite obtido no receptor, utilizando a lei dos senos (cf. Figura 18):

$$\sin z_{ipp} = \frac{R}{R+H} \sin z, \quad (109)$$

onde R é o raio médio da Terra. Existem várias funções de mapeamento mais sofisticadas.

Troposfera

Para além da ionosfera, os sinais GPS sofrem ainda os efeitos da camada electricamente neutra da atmosfera, que inclui a troposfera e a estratosfera. No entanto, dado que a troposfera contém praticamente toda a massa e todo o vapor de água, o efeito da troposfera é usado para designar o efeito total da atmosfera neutra.

A troposfera é a camada mais baixa da atmosfera. É caracterizada por um decréscimo constante da temperatura com um aumento de altitude, à média de $\sim 6.5^\circ\text{C}/\text{km}$. Esta taxa de variação varia de local para local e estação do ano. Em alguns pontos da terra, existem camadas finas (1 a 2 km) nas quais a temperatura exhibe um aumento com a altitude. Estas inversões de temperatura são comuns nas regiões árticas, por exemplo, mas também podem ocorrer em zonas temperadas.

A extensão vertical desta camada não é fixa, estando compreendida entre os 8-10 km, nas regiões polares, e os 16-19 km, nas regiões tropicais e equatoriais. O limite superior é marcado por uma descontinuidade da variação da

temperatura, a tropopausa. A Figura 20 mostra a localização aproximada da tropopausa para a América do Central e América do Norte, em Janeiro e Julho de 1992.

Contrariamente à ionosfera, a troposfera é um meio não dispersivo para frequências rádio, ou seja, o atraso de fase e de grupo são equivalentes, não sendo possível remover o efeito da troposfera usando duas frequências.

O efeito da troposfera depende do índice de refração, n . Dado que as variações do índice de refração são extremamente pequenas, introduz-se uma nova quantidade, denominada refractividade, N :

$$N = (n - 1) \cdot 10^6. \quad (110)$$

A refractividade pode ser escrita em função dos elementos meteorológicos. A expressão geral é dada por:

$$N = K_1 \left(\frac{P_d}{T} \right) Z_d^{-1} + K_2 \left(\frac{e}{T} \right) Z_w^{-1} + K_3 \left(\frac{e}{T^2} \right) Z_w^{-1}, \quad (111)$$

onde

P_d pressão parcial devida aos gases secos na atmosfera (hPa)

T temperatura (K)

e pressão parcial devida ao vapor de água (hPa)

K_i constantes a serem determinadas em laboratório

Z_d factor de compressibilidade para o ar seco

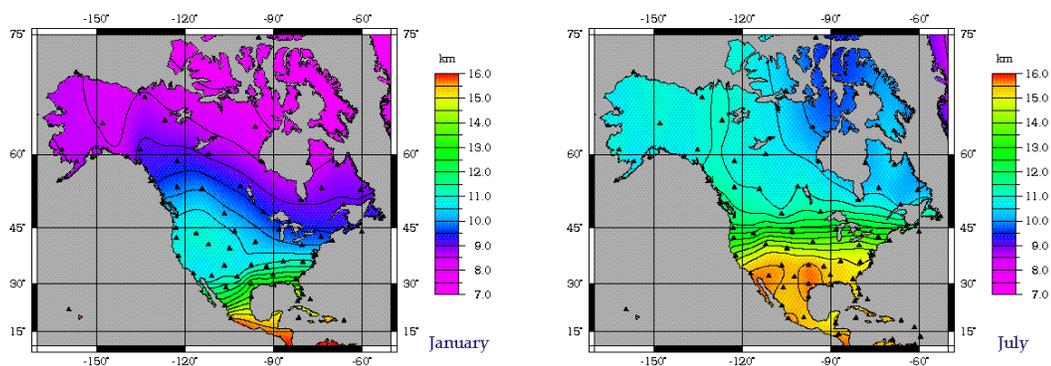


Figura 20 – Variação regional da tropopausa.

Z_w factor de compressibilidade para o vapor de água

Os factores de compressibilidade são correcções aos desvios do ar relativamente ao comportamento de um gás ideal.

Usando as constantes determinadas por Thayer [1974] e ignorando os factores de compressibilidade, obtemos:

$$N = 77.6 \left(\frac{P_d}{T} \right) + 64.8 \left(\frac{e}{T} \right) + 3.776 \times 10^5 \left(\frac{e}{T^2} \right). \quad (112)$$

O primeiro termo representa a componente seca da refractividade, enquanto os dois restantes termos representam a componente húmida, devida ao efeito do vapor de água.

O atraso troposférico sofrido por um sinal rádio na sua propagação entre o satélite e o receptor é dado, em primeira aproximação, pela seguinte expressão:

$$d_{\text{trop}} = \int_{r_s}^{r_a} [n(r) - 1] \csc \theta \, dr + \left[\int_{r_s}^{r_a} \csc \theta \, dr - \int_{r_s}^{r_a} \csc \varepsilon \, dr \right], \quad (113)$$

onde

r_a raio geocêntrico do topo da atmosfera neutra

r_s raio geocêntrico do receptor

θ ângulo de elevação aparente (ou seja, refractado)

ε ângulo de elevação geométrico (ou seja, verdadeiro)

Na Equação (113), o primeiro integral representa a diferença entre os comprimentos electromagnético e geométrico do trajecto de transmissão refractado. O segundo integral, conhecido por atraso geométrico, representa a diferença entre os trajectos refractado e rectilíneo (*ray bending*) e é significativo para ângulos de elevação inferiores a 10° .

Para um sinal na direcção do zénite do receptor, o atraso geométrico é nulo e a Equação (113) virá:

$$d_{\text{trop}}^z = \int_{r_s}^{r_a} [n(r) - 1] \, dr, \quad (114)$$

ou, usando a Equação (110):

$$d_{\text{trop}}^z = 10^{-6} \int_{r_s}^{r_a} N \, dr, \quad (115)$$

que constitui o atraso troposférico zenital.

Geralmente a componente seca e a componente devida ao vapor de água são tratadas separadamente. Usando a Equação (111), vem:

$$d_{\text{trop}}^z = 10^{-6} \int_{r_s}^{r_a} K_1 \left(\frac{P_d}{T} \right) Z_d^{-1} \, dr + 10^{-6} \int_{r_s}^{r_a} \left[K_2 \left(\frac{e}{T} \right) + K_3 \left(\frac{e}{T^2} \right) \right] Z_w^{-1} \, dr. \quad (116)$$

De um modo geral, a componente seca representa mais do que 90% do atraso troposférico total e pode ser determinada com grande precisão a partir de medições da pressão atmosférica, assumindo um bom modelo e ausência de erros na medição da pressão. Embora contribuindo com menos do que 10%, a componente húmida contribui com um maior erro, devido ao carácter variável do vapor de água na atmosfera, que torna difícil a sua modelação. Em termos quantitativos, o valor do atraso troposférico total é de cerca de 2.4 m, na direcção do zénite e ao nível do mar. Para um ângulo de elevação de 10°, este valor aumenta para cerca de 13 m.

O atraso troposférico zenital pode ser relacionado com o atraso que o sinal sofreria num dado ângulo de elevação através do uso de uma função de mapeamento. Se as funções de mapeamento são desenvolvidas separadamente para a componente seca e húmida, o atraso troposférico para um dado ângulo de elevação é dado por:

$$d_{\text{trop}}^z = d_d^z \cdot m_d(\varepsilon) + d_w^z \cdot m_w(\varepsilon) \quad (117)$$

onde

- d_d^z componente seca (hidrostática) do atraso troposférico zenital
- d_w^z componente húmida (não hidrostática) do atraso troposférico zenital
- m_d função de mapeamento para a componente seca (hidrostática)

m_w função de mapeamento para a componente húmida (não hidrostática)

Existe um grande número de modelos para a determinação do atraso troposférico zenital, bem como funções de mapeamento. Algumas das funções de mapeamento mais precisas foram desenvolvidas por Lanyi [1984], Ifadis [1986], Herring [1992] e Niell [1993]. Os três últimos modelos são baseados numa forma truncada da fracção continuada:

$$m(\varepsilon) = \frac{1}{\sin \varepsilon + \frac{a}{1 + \frac{b}{1 + \frac{c}{\sin \varepsilon + \dots}}}} \quad (118)$$

onde a , b e c são coeficientes que dependem de parâmetros meteorológicos e/ou da localização da estação. Actualmente, as Convenções do IERS 2010 recomendam a utilização das funções VMF ou GMF, desenvolvidas na Universidade Técnica de Viena.

Para além da utilização de modelos, geralmente baseado em parâmetros meteorológicos, que geralmente conduz a uma determinação pouco fiável da componente húmida, o atraso troposférico zenital pode ser determinado usando radiómetros de vapor de água (WVR - *Water Vapor Radiometer*) ou estimado juntamente com os restantes parâmetros geodésicos no processo de ajustamento, geralmente através de um processo estocástico. O WVR (Figura 21) mede a radiação emitida pelo vapor de água presente na atmosfera, permitindo a determinação da componente húmida com precisão subcentimétrica. É, no entanto, muito sensível às condições atmosféricas e pouco fiável em condições de chuva. Para além disso, é particularmente dispendioso, pelo que a sua utilização tem sido muito limitada. O atraso troposférico pode também ser determinado a partir de observações de balões meteorológicos (Figura 22).

Nos últimos anos, a estimação estocástica da componente húmida do atraso troposférico tem-se revelado como a técnica mais robusta. Na estimação de **parâmetros troposféricos**, como parte da solução por mínimos quadrados, parte-se da modelação pura como valor *a priori* para a correcção do efeito da atmosfera neutra e estima-se uma correcção a estes modelos. Nesta situação, é



Figura 21 - Radiômetro de vapor de água (WVR)

importante ter uma boa função de mapeamento, uma vez que ela aparece na matriz de configuração. Tem-se:

$$\boxed{d_{\text{trop}} = d_{\text{trop}_{\text{ap}}} + d_e^z(\tau) \cdot m_w(\varepsilon)}, \quad (119)$$

onde $d_{\text{trop}_{\text{ap}}}$ é o valor do atraso da atmosfera neutra *a priori*, m_w é a função de

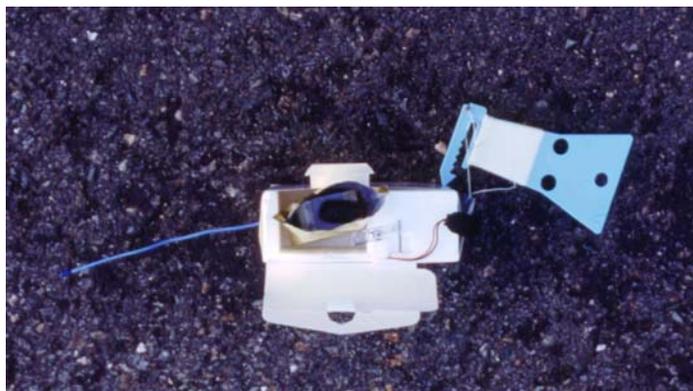


Figura 22 - Aquisição de dados de um balão meteorológico.

mapeamento a utilizar na estimação dos parâmetros troposféricos e $d_c^z(\tau)$ são os parâmetros troposféricos, estimados a cada intervalo de tempo τ . A estimação dos parâmetros troposféricos está na base das aplicações do GPS à Meteorologia.

O pressuposto da simetria esférica da atmosfera é apenas aproximadamente válido. Saastamoinen [1973] estima um erro máximo 2.0 cm para um ângulo de elevação de 10° , devido à existência de gradientes horizontais na atmosfera. Gardner [1977] indica valores muito semelhantes (emq de 3 cm, para um ângulo de elevação de 10°). Para bases muito longas (várias centenas a milhares de km), podem, por isso, ser ainda estimados **gradientes troposféricos**, que compensem a assimetria da atmosfera terrestre (segundo as direcções Norte-Sul e Este-Oeste).

A falta de simetria azimutal pode ser modelada utilizando o conceito de “atmosfera inclinada” proposto por Herring [1992]:

$$d_{\text{trop}}^{\text{az}} = m_{\text{az}}(\varepsilon) \left[\Xi_{\text{NS}} \cos \alpha + \Xi_{\text{EW}} \sin \alpha \right] \quad (120)$$

onde $d_{\text{na}}^{\text{az}}$ representa o atraso adicional provocado pelos gradientes horizontais, Ξ_{NS} e Ξ_{EW} são parâmetros que representam os atrasos devidos à assimetria azimutal, (componentes norte e este do vector gradiente), α é o azimute e $m_{\text{az}}(\varepsilon)$ é a função de mapeamento para a atmosfera inclinada, dada pela expressão seguinte:

$$m_{\text{az}}(\varepsilon) = \frac{1}{\sin \varepsilon \tan \varepsilon + 0.0032} \quad (121)$$

Com base na análise de ~3 anos de dados VLBI para um grupo de estações, Herring concluiu que a contribuição devida à assimetria azimutal da atmosfera é geralmente inferior a 10 cm, para um ângulo de elevação de 5° .

MacMillan [1995] apresenta uma expressão semelhante, mas que permite a escolha de qualquer função de mapeamento

$$d_{\text{trop}}^{\text{az}} = m(\varepsilon) \cot(\varepsilon) \left[\Xi_{\text{NS}} \cos \alpha + \Xi_{\text{EW}} \sin \alpha \right]. \quad (122)$$

PRECISÃO DO POSICIONAMENTO GPS

A precisão com que as posições podem ser determinadas com o GPS, nomeadamente em modo absoluto, depende de dois factores: geometria da configuração dos satélites e precisão da medição da distância receptor-satélite, ambos variáveis em função do tempo.

O termo usualmente utilizado para representar a precisão de uma medição receptor-satélite GPS é chamado UERE (*User Equivalent Range Error*). Este parâmetro representa o efeito combinado dos erros das efemérides, dos erros de propagação, dos erros dos relógios e do ruído do receptor. O UERE é variável, devido às variações que sofre cada um dos erros parciais, e só assume uma distribuição normal e semelhante para cada satélite da constelação a longo prazo (dias a meses), ou seja, de um modo geral, o UERE é diferente de satélite para satélite em cada instante. A contribuição do UERE resultante do Segmento Espacial e de Controlo é denominada URE (*User Range Error*) e pode ser expressa por um indicador estatístico, conhecido como URA (*User Range Accuracy*), que varia entre 0 (maior precisão) e 15.

O efeito da geometria da configuração dos satélites é expresso pelo factor de degradação de precisão (DOP – *Dilution Of Precision*), que é a razão entre a precisão do posicionamento e a precisão da medição:

$$\sigma = \text{DOP} \cdot \sigma_0 \quad (123)$$

onde σ_0 é precisão inerente à medição e σ é a precisão do posicionamento. O DOP é um escalar que representa a contribuição da geometria da configuração na precisão do posicionamento. Existem diversos factores DOP:

$\sigma_v = \text{VDOP} \cdot \sigma_0$	DOP vertical
$\sigma_H = \text{HDOP} \cdot \sigma_0$	DOP horizontal
$\sigma_P = \text{PDOP} \cdot \sigma_0$	DOP de posição (tridimensional)
$\sigma_{HT} = \text{HTDOP} \cdot \sigma_0$	DOP horizontal e tempo
$\sigma_G = \text{GDOP} \cdot \sigma_0$	DOP de posição (3D) e tempo

O factor DOP varia obviamente com o tempo (devido ao movimento dos satélites) e com a localização geográfica e deverá ser tão pequeno quanto possível, como facilmente se depreende da Equação (123).

O GDOP constitui um bom critério na selecção dos quatro melhores satélites visíveis. Existe uma grande correlação entre o volume do tetraedro formado pelos vectores que ligam o receptor aos diferentes satélites: um GDOP elevado está associado a uma geometria de configuração em que os satélites estão muito juntos; um GDOP muito baixo está geralmente associado a uma constelação com um dos satélites sobre o observador e os outros 3 satélites igualmente espaçados em azimute no horizonte.

Sendo V o volume do tetraedro, tem-se:

$$\text{GDOP} \cong \frac{1}{V} \quad (124)$$

O volume é máximo quando um dos 4 satélites está na direcção do zénite do observador e os outros 3 estão separados 120 graus entre si e possuem, em relação ao observador, ângulos de elevação baixos.

O valor dos DOPs pode ser calculado a partir da raiz quadrada do traço da matriz de variâncias-covariâncias dos parâmetros, que por sua vez é definida como:

$$\mathbf{C}_{\hat{\mathbf{x}}} = \sigma_0^2 (\mathbf{A}^T \mathbf{A})^{-1}, \quad (125)$$

ou seja,

$$\mathbf{C}_{\hat{\mathbf{x}}} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix}, \quad (126)$$

assumindo uma matriz de pesos unitária. Os elementos da diagonal desta matriz são usadas na definição dos diversos DOPs, por exemplo,

$$\text{GDOP} = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2}{\sigma_0^2}} \quad (127)$$

$$\text{PDOP} = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\sigma_0^2}} \quad (128)$$

$$\text{HDOP} = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{\sigma_0^2}}. \quad (129)$$

A divisão pelo factor de variância *a priori* retira a informação relativa à precisão das observações, dado que apenas estamos interessados na geometria do problema. As seguintes relações são imediatas:

$$\text{GDOP}^2 = \text{PDOP}^2 + \text{TDOP}^2 \quad (130)$$

$$\text{GDOP}^3 = \text{HDOP}^3 + \text{VDOP}^3 + \text{TDOP}^3. \quad (131)$$

A geometria dos satélites visíveis é um factor importante na obtenção de resultados de boa qualidade, especialmente para posicionamento do ponto isolado e aplicações cinemáticas, que geralmente exigem baixos valores de PDOP (inferior a 5-6) e um número mínimo de satélites visíveis e, por isso, os valores DOP são bastante úteis no planeamento de períodos de observação (com a actual constelação de satélites, tal não é problemático). A Figura 23 mostra o GDOP e o número de satélites visíveis para a região de Lisboa.

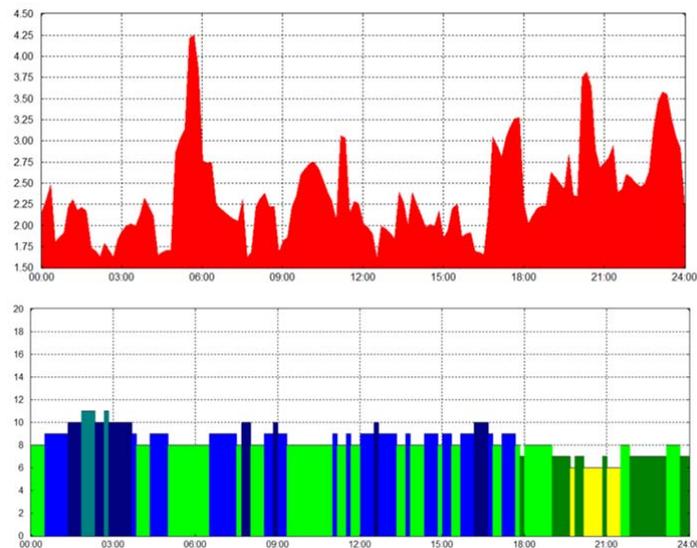


Figura 23 – Valor de GDOP (gráfico superior) e número de satélites visíveis (gráfico inferior) para região de Lisboa (simulação para o dia 25 de Maio de 2011).

RESOLUÇÃO DAS AMBIGUIDADES DE CICLO

Tal como já foi referido, as observações de fase em GPS são muito mais precisas do que as observações de pseudo-distância e conduzem, por isso, a um melhor posicionamento. No entanto, as medições de fase são ambíguas, dado que se desconhece o número inteiro de ciclos decorridos entre a época de transmissão e a época de início de registo, a **ambiguidade de ciclo**. A determinação do valor inteiro da ambiguidade de ciclo para cada satélite constitui uma extensão do processo de ajustamento e denomina-se **resolução das ambiguidades** (ou fixação de ambiguidades). Embora a resolução das ambiguidades não seja um factor essencial em posicionamento estático com sessões de observações longas e boa geometria dos satélites, existe no entanto, em geral, uma melhoria na exactidão e precisão dos vectores estimados em soluções com ambiguidades fixas relativamente às **soluções flutuantes** (ambiguidades não resolvidas). Para sessões muito curtas e modos de posicionamento cinemático, a resolução das ambiguidades é um factor essencial e conduziu ao desenvolvimento de um conjunto de técnicas de resolução de ambiguidades, que diferem essencialmente no conjunto de constrangimentos específicos que adoptam. Os quatro constrangimentos geralmente utilizados são a geometria instantânea dos satélites visíveis, a variação temporal da geometria dos satélites, bases curtas e ambiguidades inteiras.

As técnicas de resolução das ambiguidades podem classificar-se em diferentes classes: (1) técnicas apoiadas em modos operacionais especiais; (2) técnicas de procura de ambiguidades em domínios (observações, coordenadas, ou ambiguidades estimadas); (3) técnicas de recuperação de ambiguidades; (4) técnicas integradas.

As técnicas de resolução de ambiguidades apoiadas em modos operacionais especiais exigem que o posicionamento se faça segundo regras específicas. A técnica mais conhecida desta classe é a troca de antenas, que consiste na colocação de dois receptores em dois pontos distintos em observação simultânea. Após serem obtidas as primeiras épocas de observações, procede-se à troca de antenas (a antena do receptor A passa a ser estacionada na marca correspondente ao receptor B e vice-versa), mantendo a sintonia dos satélites durante esta mudança. Recolhem-se mais algumas observações para esta nova configuração e regressa-se à configuração inicial, para serem obtidas mais algumas épocas de observação. Se esta operação for realizada sob os constrangimentos de base curta e boa geometria instantânea de satélites, o

vector entre os dois pontos pode ser determinado com precisão milimétrica e permitir uma fácil resolução de ambiguidades. Para tal, consideramos que a contribuição da atmosfera (ionosfera e troposfera) é nula. Se os pontos tiverem coordenadas conhecidas, as diferenças duplas de fase correspondem às diferenças duplas de ambiguidades. Esta técnica era utilizada para inicialização das ambiguidades inteiras nos levantamentos pára-avança e cinemáticos (ver Secção Métodos de Posicionamento GPS). Hoje em dia, como resultado de uma melhoria significativa dos receptores e boas configurações, com vários satélites visíveis, as ambiguidades podem ser resolvidas instantaneamente e sem necessidade de inicialização. Estas técnicas pertencem a uma classe particular denominada resolução instantânea de ambiguidades ou OTF (*On-The-Fly*).

De um modo geral, as técnicas de busca de ambiguidades em domínios apresentam uma série de características comuns, que compreendem diversas fases: obter de uma solução inicial; estabelecer um domínio de busca; estabelecer uma metodologia para pré-seleccionar conjuntos de ambiguidades ou posições dentro desse domínio a serem testados; geração de soluções fixas com os conjuntos candidatos à solução final; aceitação ou rejeição da solução obtida, utilizando um determinado critério. Esta classe apresenta 3 variantes principais: observações, coordenadas e ambiguidades.

Na **busca de ambiguidades no domínio das observações**, a resolução das ambiguidades é feita directamente a partir das observações de pseudo-distância e fase.

Se tivermos apenas uma frequência, uma das formas de determinar as ambiguidades consiste em estimar estes parâmetros em conjunto com os restantes parâmetros de interesse, no processo de ajustamento (com parâmetros inúteis). Se o valor real estimado para as ambiguidades for “próximo” de um valor inteiro, podemos então assumir que o valor dessas ambiguidades correspondem ao valor arredondado mais próximo e fixar essas ambiguidades num ajustamento posterior. Se não for possível fixar todas as ambiguidades em simultâneo, podemos utilizar um processo sequencial, isto é, fixam-se as ambiguidades possíveis após o primeiro ajustamento, deixando as restantes como parâmetros ainda a estimar, e efectua-se um novo ajustamento, que tem neste caso um maior número de graus de liberdade, e que poderá, por isso, conduzir a uma melhor determinação das ambiguidades restantes. De um modo geral, este processo funciona bem para bases curtas e sessões superiores a cerca de 1 hora, utilizando diferenças duplas. De notar que os restantes erros não modelados (por exemplo, a ionosfera) podem influenciar de forma significativa

a estimação correcta das ambiguidades. O método é também sensível à influência do multitrajeto.

Se operarmos com receptores de dupla frequência, a resolução das ambiguidades pode ser facilitada com a utilização de combinações lineares de observações, tal como a combinação de banda larga e de banda estreita.

Em algumas situações, podemos utilizar as medições de pseudo-distância para determinar as ambiguidades (dado serem medições não ambíguas) e derivar posteriormente as ambiguidades da fase (solução *pseudo-distância amaciada com fase*). As restrições à aplicação desta técnica são as bases curtas e ambiguidades inteiras. A combinação de pseudo-distâncias com fase conduz a uma solução que é independente da geometria (técnica geometricamente livre), que pode ser aplicada a bases médias e longas, mas requer observações de pseudo-distância com boa qualidade.

Se as componentes da base forem conhecidas com precisão centimétrica, as ambiguidades inteiras podem ser estimadas no domínio das coordenadas. Este método utiliza um procedimento de busca numa grelha, por tentativa e erro. Numa primeira fase, eliminam-se as ambiguidades usando diferenças triplas, de modo a obter uma primeira estimativa da posição e respectiva precisão; constrói-se um domínio de busca (um paralelepípedo centrado na posição estimada e lados relacionados com a incerteza dessa estimativa, através da aplicação de um factor arbitrário); cada um dos pontos da grelha tridimensional (que representa um posição ligeiramente diferente da estimada) é candidato à solução, pelo que se calculam as ambiguidades, mantendo a posição candidata fixa; são seleccionadas as soluções que conduzem a valores de ambiguidades que se aproximam o mais possível de números inteiros; uma vez seleccionada “a melhor” solução, fixam-se as ambiguidades e calcula-se a posição final. O método da função de ambiguidade (AFM – *Ambiguity Function Method*) baseia-se nesta técnica. Os métodos baseados neste conceito são, em termos computacionais, pouco eficientes e pouco adaptados para as aplicações em tempo útil.

O procedimento de busca no domínio das ambiguidades começa por obter uma solução flutuante. Seguidamente inicia um processo de busca do conjunto de ambiguidades que minimiza a forma quadrática dos resíduos das observações. Fazem parte deste grupo, o processo rápido de resolução de ambiguidades (FARA – *Fast Ambiguity Resolution Approach*), a técnica de busca de ambiguidades por mínimos quadrados (LSAST – *Least-Squares Ambiguity*

Search Technique), o método de decomposição de Cholesky, o filtro de busca rápida de ambiguidades (FASF – *Fast Ambiguity Search Filter*) e o método de decomposição. O número de operações a realizar neste tipo de métodos é extremamente elevado, pelo que a aplicação de alguns critérios de otimização são essenciais. Se forem observados n_{SVs} satélites e se W for a janela de busca (número de ciclos no interior do intervalo de busca), o número de candidatos é dado por $(2W+1)^{n-1}$. Como exemplo, se W for 10 (representando uma incerteza da ordem dos 2 m) e forem observados 6 satélites, o número de candidatos é de $\sim 4.1 \times 10^6$. Se tivermos 10 satélites, este número aumenta para $\sim 7.9 \times 10^{11}$.

O ajustamento de decorrelação de ambiguidades por mínimos quadrados (LAMBDA – *Least-Squares Ambiguity Decorrelation Adjustment*) também se pode incluir neste grupo. A aplicação deste método envolve todos os constrangimentos e requer a utilização de receptores de fase e pseudo-distância. A filosofia do método consiste na transformação de ambiguidades fortemente correlacionadas em ambiguidades pouco correlacionadas, que tem como benefício imediato a redução significativa do número de candidatos, um dos principais problemas com as técnicas de busca em domínios. O método faz inicialmente uma decorrelação das diferenças duplas de ambiguidades para depois determinar essas ambiguidades no domínio dos inteiros, funcionando como uma extensão do método dos mínimos quadrados ao domínio dos inteiros.

A técnica de recuperação de ambiguidades consiste na reestimação das ambiguidades em situações de ocorrência de saltos de ciclo, sem perda de dados significativa (inferior a poucos minutos). Requer a inicialização das ambiguidades no início da sessão e receptores de fase e pseudo-distância de boa qualidade.

Finalmente, as técnicas integradas combinam os procedimentos dos diferentes métodos de busca, numa tentativa de encontrar simultaneamente o valor máximo da função de ambiguidade e a minimização da forma quadrática dos resíduos. Incluem-se ainda dentro desta classe aquelas técnicas que recorrem a dados provenientes de outros sensores, tais como os sistemas de navegação inercial.

Com a introdução de uma nova frequência, os métodos para a resolução das ambiguidades têm aumentado de forma significativa, havendo um novo conjunto de métodos disponíveis na literatura.

MÉTODOS DE POSICIONAMENTO GPS

Nos últimos anos, o posicionamento com GPS é feito recorrendo a vários métodos de observação. Este capítulo descreve de forma sucinta as diferentes metodologias usadas em modo estático e cinemático com o GPS, usando a fase como observável principal. Alguns destes métodos já não são usados actualmente.

A distinção inicial entre posicionamento estático e cinemático era feita de acordo com o modo de resolução das ambiguidades. No posicionamento estático, as ambiguidades de ciclo eram consideradas como parâmetros desconhecidos, que seriam determinadas no processo de ajustamento, em conjunto com as coordenadas das estações e outros parâmetros de interesse. No posicionamento cinemático, estas ambiguidades eram determinadas usando procedimentos apropriados antes de se dar início ao posicionamento propriamente dito.

No posicionamento **estático** relativo, os receptores utilizados ocupam posições fixas. O tempo necessário de estacionamento dos receptores depende de vários factores, tais como o comprimento da base, número de satélites observados, geometria entre os receptores e os satélites, medições em frequência simples ou dupla, precisão final pretendida, etc. Embora não seja obrigatório, um mínimo de quatro satélites deverão ser visíveis durante pelo menos parte da sessão de observação.

As incógnitas principais neste tipo de posicionamento são as diferenças de coordenadas entre as antenas dos receptores GPS, as ambiguidades de ciclo e, caso não se usem técnicas diferenciais, o desfasamento entre os relógios dos receptores. A resolução das ambiguidades, em particular, exige tempos de observação particularmente longos (dependentes do comprimento da base), que permita uma variação na geometria receptores-satélites. As sessões longas permitem, para além disso, uma redução significativa dos efeitos de multitrajecto e ruído do receptor.

O posicionamento **rápido-estático** (ou estático-rápido) corresponde a uma sessão estática de curta duração (5 a 10 minutos). Para este tipo de posicionamento, um receptor pára numa dada estação durante 5 a 10 minutos, e a sua localização é referenciada a um receptor operando continuamente, a uma distância que pode ir até cerca de 20 km. O principal problema neste tipo de posicionamento é o da resolução das ambiguidades de ciclo.

Um dos procedimentos possíveis consiste em utilizar pseudo-distâncias muito precisas, que permite resolver as ambiguidades muito mais rapidamente (pseudo-distância amaciada com fase). Este método requer medições de fases e pseudo-distâncias, em ambas as frequências. Outro processo, que não requer pseudo-distâncias, usa um algoritmo de resolução rápida de ambiguidades (OTF). Este procedimento dá bons resultados para bases pequenas, se possuímos medições de fase em ambas as frequências, e não requer que o receptor se mantenha ligado durante o trajecto entre os diversos pontos.

O posicionamento **pseudo-cinemático** (também conhecido como pseudo-estático, reocupação, estático instantâneo, ou estático intermitente) é idêntico ao método rápido-estático, mas requer que cada ponto seja estacionado uma segunda vez, após um intervalo de tempo significativo (por exemplo, uma hora), que permita uma geometria de observação diferente. Com este procedimento pretende-se ligar a fase entre as duas sessões, o que é idêntico a um posicionamento estático de longa duração, com uma grande lacuna de observações. Durante este intervalo entre sessões, outros pontos podem ser estacionados. Desta forma, este tipo de posicionamento é mais económico do que o estático convencional.

No posicionamento **pára-avança** (*Stop and Go*), os valores das ambiguidades de ciclo são determinados antes de se iniciar o posicionamento, através de procedimentos adequados. Após resolvidas as ambiguidades, os únicos parâmetros que permanecem desconhecidos são as diferenças de coordenadas entre as antenas dos receptores. Deste modo, este posicionamento consiste apenas no transporte de um receptor a todos os pontos da rede a observar, com breves períodos de observação em cada um deles. Em princípio, uma época é suficiente para obter instantaneamente as coordenadas de cada estação envolvida, se forem recebidos sinais de pelo menos 4 satélites, tornando este procedimento vantajoso em termos económicos. Uma vez que o requisito básico é a pré-determinação das ambiguidades, o transporte do receptor de um ponto para outro deverá ser feito muito cuidadosamente, de forma a evitar obstrução de sinal e saltos de ciclo (contrariamente ao método rápido-estático e pseudo-cinemático).

O posicionamento **cinemático** não requer que o receptor móvel fique em modo estático em qualquer período de tempo, antes, durante ou após a sessão. Desta forma, a resolução de ambiguidades é feita com base nas medições obtidas durante a sessão (OTF). A resolução de ambiguidades pode ser feita usando apenas informação da fase, se forem feitas observações contínuas para pelo

menos quatro satélites, sem perda de sintonia (ou seja, sem saltos de ciclo). No entanto, a utilização da informação de pseudo-distâncias é também usada frequentemente na resolução das ambiguidades.

As aplicações em tempo-real necessitam não só de um receptor a funcionar como referência (receptor base) e outro que se vai deslocando de ponto para ponto (receptor móvel, também designado como *rover*), mas também uma ligação rádio apropriada, necessária para enviar correcções da estação base para a estação móvel. Estas técnicas de **GPS Diferencial** (DGPS) existem desde meados de 1980, para correcções a pseudo-distâncias (essencialmente utilizadas em navegação marítima e aérea), com precisões de cerca de um metro. Actualmente, existem também novos serviços baseados em correcções diferenciais às pseudo-distâncias, designados por SBAS (*Satellite-Based Augmentation System*) ou GBAS (*Ground-Based Augmentation System*).

Mais recentemente esta técnica de GPS diferencial foi adaptada para a utilização das observações de fase. Estes sistemas, são conhecidos como **RTK** (acrónimo de *Real Time Kinematic*), proporcionam precisão centimétrica em tempo real. Os métodos baseados em RTK (quer em modo-base, quer em modo-rede) constituem, a par com o posicionamento relativo, os processos mais frequentes na utilização do GPS em Topografia e serão abordados com mais detalhe. Nos anos mais recentes, o posicionamento pontual preciso (PPP) tem tido grande desenvolvimento, em todo o tipo de aplicações. A Figura 24 representa a visão mais clássica dos métodos de posicionamento GPS.

Posicionamento Pontual Preciso (PPP)

O método de posicionamento pontual preciso (PPP - *Precise Point Positioning*) foi desenvolvido no final da década de 1990, inicialmente para aplicações estáticas, mas tornou-se uma técnica de posicionamento com grande potencial, rapidamente explorado para incluir também as aplicações cinemáticas. O PPP utiliza combinação de pseudo-distâncias e fase, sendo um posicionamento do tipo absoluto. Nesse sentido, o método requer a modelação de todos os erros presentes na equação de fase de dados não diferenciados, com particular realce para os erros dos relógios e das órbitas. As aplicações de precisão com esta técnica só é possível graças às baixas incertezas dos erros dos relógios satélites e das órbitas dos satélites, determinadas por instituições como o IGS. A principal vantagem da técnica PPP é de poder utilizar apenas um receptor. Nas situações de utilização como sensor de movimento, por exemplo, essa particularidade é especialmente vantajosa. O referencial associado às coordenadas PPP é sempre o correspondente ao utilizado na geração das órbitas.

Para além dos erros comuns aos outros métodos de posicionamento (relógios, órbitas, efeitos atmosféricos, multitrajecto, etc.), a técnica implica a correcção de outros erros: marés, carga oceânica, *wind-up* da fase, e atraso diferencial de grupo dos satélites. As marés incluem as **marés terrestres** (deformação elástica da Terra) e as **marés oceânicas** (que são influenciadas pela geografia da costa e pela topografia oceânica). A **carga oceânica** é provocada pela flutuação do nível do mar, devido às marés, com efeito mais pronunciado nas proximidades da costa. O **wind-up de fase** é devido à variação de fase medida em consequência da variação de orientação das antenas. O **atraso diferencial de grupo** dos satélites corresponde ao enviesamento instrumental L1-L2, variável de satélite para satélite, mas enviado na mensagem de navegação.

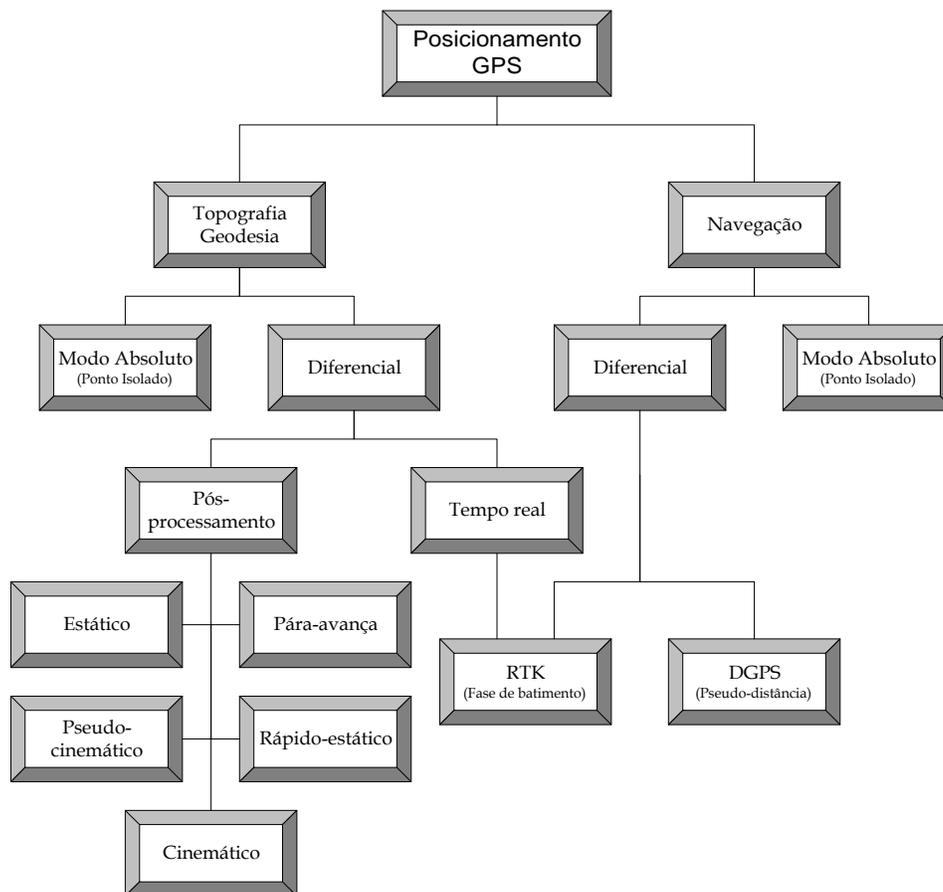


Figura 24 - Classificação "clássica" dos métodos de posicionamento GPS

A técnica PPP tem níveis de incerteza próximos do posicionamento relativo (da ordem de cm), mas apenas após algumas horas de observação.

Sistemas de aumento

Os sistemas de aumento (o termo *aumentação* é também utilizado) com base em satélites ou meios terrestres (SBAS – *Satellite-Based Augmentation Systems* ou GBAS – *Ground-Based Augmentation Systems*) surgem com a necessidade de responder a diferentes questões:

- **Exactidão** – grau de conformidade do valor calculado com o valor verdadeiro
- **Integridade** – sistema de alerta para situações em que o sistema não deve ser usado
- **Continuidade** – capacidade do sistema realizar as suas funções sem interrupção durante a operação pretendida
- **Disponibilidade** – percentagem de tempo em que os serviços do sistema estão disponíveis

Alguns dos SBAS desenvolvidos ou em vias de desenvolvimento incluem:

- DGPS (âmbito nacional) – Differential GPS
- WAAS – *Wide Area Augmentation System* (EUA)
- LAAS – *Local Area Augmentation System* (EUA)
- EGNOS – *European Geostationary Navigation Overlay System* (Europa)
- MSAS – *Multi-functional Satellite-based Augmentation System* (Japão)
- SNAS – *Satellite Navigation Augmentation System* (China)

NDGPS (*Nationwide Differential GPS*), EUA

Este sistema proporciona, em média, uma exactidão melhor que 10 m, com valores que se aproximam de 1 m, na vizinhança do radiofarol. A degradação espacial em exactidão esperada para um sistema deste tipo é da ordem de 1 m por 150 km de afastamento em relação ao radiofarol emissor de correcções.

DGPS em Portugal

O Instituto Hidrográfico implementou em Portugal uma rede DGPS. As estações desta rede são Sagres, Cabo Carvoeiro, Horta e Porto Santo (complementadas por Finisterra e Rota, em Espanha). Os valores de exactidão proporcionados por testes efectuados pelo IH revelam um desempenho inferior ao registado pela rede NDGPS.

WAAS

O WAAS foi concebido para a navegação aérea e é da responsabilidade da FAA (*Federal Aviation Administration*). É utilizável nas fases de voo em rota e em aproximação de precisão (Categoria I – exactidão de posição vertical de 4.0 m). Inclui 3 serviços: (1) integridade dos dados GPS e de satélites geostacionários (GEO); (2) correcções diferenciais para GPS e GEO; (3) disponibilidade e continuidade elevadas.

LAAS

O LAAS é um sistema semelhante ao WAAS, mas de carácter local (GBAS). O LAAS utiliza difusão de correcções diferenciais via VHF, sendo utilizável nas fases de voo em rota e em aproximação de precisão, envolvendo categorias I, II e III (as categorias II e III exigem exactidão de posição vertical de 2.5 m).

Posicionamento cinemático em tempo real

O posicionamento GPS cinemático em tempo real (RTK – *real-time kinematic*) foi desenvolvido em meados da década de 1990. Este método de posicionamento envolve uma estação de referência, localizada num ponto de coordenadas conhecidas, e uma estação móvel (ou várias estações móveis). A estação de referência transmite, em tempo útil (“tempo-real”), medições brutas ou correcções às observações para o receptor móvel, usando um ligação de comunicação de dados adequada:

- Ligação rádio (VHF, UHF, FM, sinal TV, ...)
- Telemóvel (GSM – *Global System for Mobile communications*, GPRS – *General Packet Radio Service*)
- Satélite
- Internet

Este método permite reduzir o efeito de erros atmosféricos (ionosfera e atmosfera neutra) e erros orbitais, permitindo precisão centimétrica no posicionamento. A utilização do RTK depende do alcance do modo de transmissão de dados utilizado e do facto dos erros atmosféricos e orbitais sofrerem descorrelação espacial (uma vez que os erros se descorrelacionam com o aumento da distância ao receptor base, deixam de poder ser eliminados com a diferenciação das observações). Assim, a utilização do RTK não deverá exceder os 20 km, para permitir uma resolução de ambiguidades rápida e fiável, sendo esta questão fundamental em RTK.

Redes-RTK

Os erros já referidos podem ser mais eficazmente modelados se, em vez de utilizar apenas uma estação de referência (**base-RTK**), se utilizar uma rede de estações de referência em torno da estação móvel. Este é o conceito das redes GPS RTK (**rede-RTK**). O conceito rede-RTK apresenta também vantagens económicas, pois permite reduzir o número de equipamentos e pessoal no campo e otimizar o tempo dispendido nas operações de levantamento (não há necessidade de deslocar as estações de referência, quando a distância a esta estação aumenta). As vantagens das redes RTK podem ser resumidas aos seguintes pontos essenciais:

- Maior espaçamento entre estações (até 100 km, sendo o tempo necessário à resolução das ambiguidades o principal factor constrangedor da distância máxima)
- Maior produtividade
- Uniformidade de coordenadas
- Maior fiabilidade na resolução de ambiguidades, logo posições calculadas mais fiáveis e com maior exactidão

O posicionamento recorrendo a redes RTK envolve várias componentes: 1) a geração das correcções; (2) a interpolação das correcções; (3) a transmissão das correcções.

Geração de correcções

Tal como no conceito de base-RTK, a resolução das ambiguidades em tempo-real é absolutamente necessária nas redes RTK e constitui um problema, devido às maiores distâncias entre as estações de referência (várias dezenas de km). Embora as coordenadas das estações sejam conhecidas com precisão, os erros de observação associados a uma modelação deficiente de erros conduzem a esta dificuldade. As ambiguidades devem ser resolvidas para um nível comum (“todas” as ambiguidades fixas relativamente a um mesmo satélite de referência). A resolução das ambiguidades para a rede é um passo fundamental e precede a geração das correcções de rede.

Interpolação das correcções

Depois de calculadas as correcções para as diferentes estações da rede RTK, elas têm que ser interpoladas para a posição aproximada do receptor móvel, utilizando um algoritmo apropriado. São conhecidos vários algoritmos (ver Dai *et al.* [2000]) que permitam modelar ou interpolar os erros dependentes da distância entre as estações de referência e as estações móveis (como a

interpolação linear ou interpolação em grelha, entre outros). Os erros ionosféricos e orbitais devem ser modelados individualmente para cada satélite e a correções troposféricas devem ser estimadas para cada estação. A qualidade das correções aumenta com o número de estações de referência que são utilizadas. As variações ionosféricas têm uma maior variação temporal, pelo que as correções têm que ser enviadas ao receptor móvel com maior frequência.

Transmissão das correções

O processo é concluído com a transmissão das correções de rede para o receptor móvel. As técnicas de transmissão desenvolvidas ao longo do tempo podem dividir-se em dois grandes grupos, de acordo com o tipo de comunicação utilizado: unidireccional e bidireccional.

No tipo de **comunicação unidireccional**, uma ligação de dados transmite continuamente as correções de rede (ou coeficientes para aplicação a um modelo) ao utilizador. O conceito FKP (*Flächen-Korrektur-Parameter*) é um exemplo de aplicação deste tipo de comunicação.

Na **comunicação bidireccional**, o utilizador e o centro de processamento de dados da rede estabelecem uma ligação entre eles, que funciona em ambos os sentidos. O conceito VRS (*Virtual Reference Station*) é do tipo bidireccional.

Formato RTCM

A utilização das redes RTK obriga os serviços responsáveis a gerar correções para receptores que operam na zona de influência da rede, utilizando um formato standardizado. Esta padronização da informação de rede é também necessária para reduzir os tamanhos das mensagens de correção das redes RTK. As mensagens em “tempo-real” que permitem uma adequada interoperacionalidade entre diferentes fabricantes de equipamentos obedecem a um padrão, criado pelo RTCM SC-104 (*Special Committee 104 of The Radio Technical Commission for Maritime Services*, que aborda o tópico *Differential GNSS*). As diferentes versões de RTCM são:

- RTCM 1.0 – introduzida em 1985
- RTCM 2.0 – introduzida em 1990
- RTCM 2.1 – introduzida em 1994 (inclui pela primeira vez, dados de fase)
- RTCM 2.2 – introduzida em 1998 (inclui pela primeira vez, dados de fase para outros sistemas de navegação global por satélite (e.g. GLONASS))
- RTCM 2.3 - introduzida em 2001 (inclui variações de centro de fase da antena)

- RTCM 3.0 – introduzida em 2005/2006, inclui dados para redes RTK

A versão 3.0 das mensagens RTCM permite uma melhor flexibilidade e eficiência.

As mensagens mais importantes na versão 2.3 são apresentadas na Tabela 2.

Tabela 2 – Mensagens RTCM, versão 2.3

Tipo	Descrição
1	Correcções DGPS
2	Correcções DGPS
3	Parâmetros da estação de referência
18	Dados brutos (pseudo-distância)
19	Dados brutos (fase)
20	Correcções diferenciais (pseudo-distância)
21	Correcções diferenciais (fase)

A última versão RTCM é designada RTCM 10403.1 (**Differential GNSS (Global Navigation Satellite Systems) Services - Version 3**). Esta versão, que corresponderia à versão 3.1, incorpora correcções de rede-RTK e novas mensagens para GPS e GLONASS.

Nas redes RTK, a correlação espacial e temporal dos erros que dependem da distância são determinados usando observações de várias estações de referência GPS, com coordenadas conhecidas com precisão. Os resíduos das observações para cada uma destas estações são depois utilizados para obter correcções diferenciais para qualquer receptor a operar na zona de influência da rede. Para o cálculo e difusão destas correcções existem essencialmente três conceitos, denominados VRS, FKP e MAC. Os conceitos distinguem-se fundamentalmente no tratamento matemático dos parâmetros nas equações de observação.

VRS - Virtual Reference Station

O método VRS utiliza comunicações bidireccionais, pelo que o receptor móvel comunica com o centro de processamento de dados (CPD). O utilizador

transmite a sua posição aproximada para o CPD, utilizando o protocolo NMEA (*National Marine Electronics Association*). O CPD gera observações de fase e pseudo-distância para a posição aproximada que recebeu, utilizando modelos de correções, e as observações virtuais são enviadas ao utilizador, usando o protocolo RTCM, por exemplo, criando assim uma estação virtual. É então estabelecida uma base-RTK entre a estação virtual e o utilizador que, devido à curta distância que é estabelecida, permitirá resolver as ambiguidades de forma fácil e obter uma posição consistente com a rede RTK. O VRS tem como **vantagens**: (1) permitir uma modelação mais complexa dos erros atmosféricos, uma vez que usa a informação de toda a rede; (2) embora necessite comunicações bidireccionais, estas são preferíveis, uma vez que têm a capacidade de transmitir informação adicional e, em geral, as infraestruturas para GSM e GPRS já existem. As principais **desvantagens**: (1) requer comunicações bidireccionais, uma vez que o receptor móvel tem que transmitir a sua posição; (2) quando o receptor móvel está em modo cinemático, tem uma variação contínua das suas coordenadas, dificultando a aplicação de um modelo de correções. É um método com grande aplicação em redes RTK.

FKP - Flächen-Korrektur-Parameter

O método FKP usa a informação das estações de referência da rede de forma a calcular parâmetros lineares que descrevem o efeito dos erros atmosféricos e das órbitas. Estes parâmetros são depois difundidos para os utilizadores sob a forma de gradientes, que são utilizados num processo de interpolação dos erros da rede para a posição específica do utilizador (parametrisação polinomial). O utilizador recebe os parâmetros da estação de referência mais próxima. O FKP tem a **vantagem** de permitir uma maior flexibilidade na escolha do meio de difusão das correções, dada a natureza unidireccional das comunicações. As principais **desvantagens** do método são: (1) a limitação na modelação de erros, uma vez que os modelos de correção são mais simples; (2) a degradação da interpolação, quando as redes são mal configuradas (espaçamento entre estações grande ou desigual).

MAC - Master Auxiliary Concept

A ideia básica do Conceito Principal-Auxiliar (MAC) é utilizar um estação de referência principal e os dados de observação a ela associados e informação reduzida de outras estações de referência auxiliares, localizadas na vizinhança da zona de operação do receptor móvel. Desta forma, o volume de informação das correções de rede a enviar é reduzido.

O MAC usa as diferenças de correcções dispersivas (ionosfera) e não dispersivas (e.g. troposfera) para comprimir a informação de rede RTK sem necessidade de utilizar modelos de correcção padrão. Dado que a correcção da componente dispersiva é cerca de uma ordem de magnitude superior à componente não dispersiva, elas podem ser transmitidas com taxas de actualização diferentes. Os seguintes valores de actualização são consensuais: 120 s para correcções orbitais, 30 s para correcções troposféricas e 10 s para correcções ionosféricas.

Uma das grandes **vantagens** do MAC é a flexibilidade que a unidade móvel tem para efectuar uma interpolação simples, eficiente e robusta das correcções de rede ou aplicar um algoritmo de cálculo mais complexo, uma vez que dispõe de toda a informação. O método usa comunicações unidireccionais.

Uma variante do método MAC é o **i-MAX**, que consiste em correcções individualizadas. O método i-MAX requer comunicações bidireccionais, à semelhança do VRS. Neste método, o CPD escolhe as estações de referência mais apropriadas e a estação mais próxima do utilizador é escolhida como estação principal, ou seja, não é gerada nenhuma estação virtual, como no VRS. O i-MAX foi desenvolvido pela marca Leica, de forma a permitir que os receptores mais antigos incapazes de interpretar as mensagens da versão 3 do protocolo RTCM, pudessem efectuar correcções de rede-RTK.